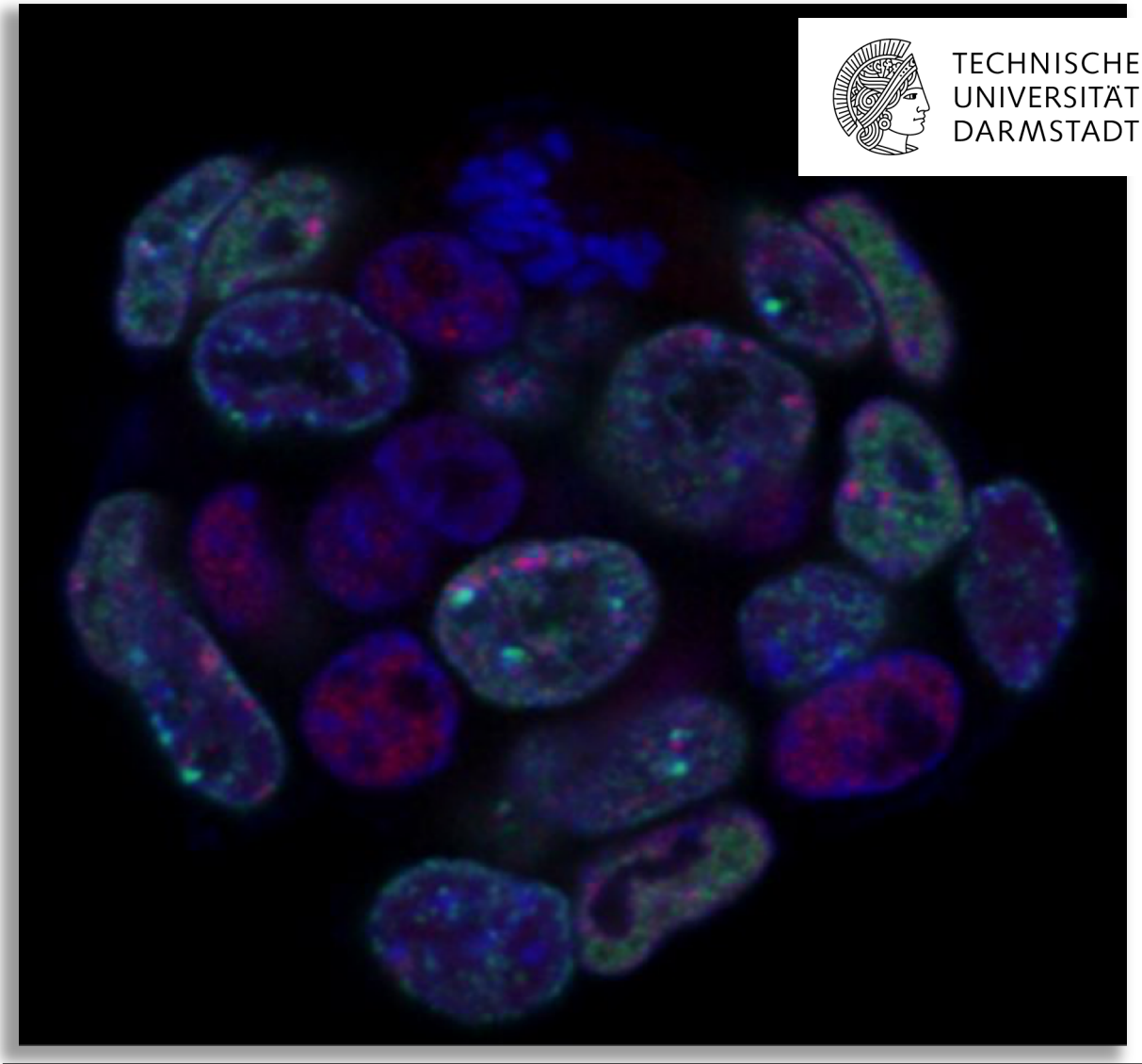
DNA replication dynamics in embryonic stem cells

DNA Replikationsdynamiken in embryonalen Stammzellen
Dem Fachbereich Biologie der Technischen Universität Darmstadt zur Erlangung des akademischen Grades Doctor rerum naturalium vorgelegte Dissertation von

Patrick Weber aus Mannheim

1. Referent/Referentin: Prof. Dr. M. Cristina Cardoso
2. Referent/Referentin: Prof. Dr. Bodo Laube



TECHNISCHE UNIVERSITÄT DARMSTADT

DNA replication dynamics in embryonic stem cells

vom Fachbereich Biologie
der Technischen Universität Darmstadt

zur Erlangung des Grades
Doctor rerum naturalium
(Dr. rer. nat.)

Dissertation

von Patrick Weber
aus Mannheim

Erstgutachter/in: Prof. Dr. M. Cristina Cardoso
Zweitgutachteri/in: Prof. Dr. Bodo Laube

Tag der Einreichung: 19. September 2018

Tag der mündlichen Prüfung: 09. November 2018

Darmstadt 2018

Es endet so, wie es begonnen hat

SUMMARY

DNA replication is one of the most fundamental processes in every living organism, required for the propagation of a cells (epi)genetic information. To ensure the error-free duplication and transmission of all genetic material exactly once per cell cycle, DNA replication follows a tightly controlled spatio-temporal program which is conserved across many species. Especially in higher eukaryotic organisms, however, the regulation of this program needs to be dynamic and flexible enough, in order to allow the coordination with other DNA-dependent processes like transcription and DNA repair. Furthermore, early studies already showed that the replication timing program undergoes remarkable changes during cellular development that can be controlled on multiple levels. During the early embryonic stages of *Drosophila* and *Xenopus*, for example, DNA duplication is achieved in the order of minutes and depends on the availability of maternal factors that allow rapid DNA synthesis in the absence of transcription, a specialized cell cycle organization characterized by the lack of gap phases which allows cells to oscillate between S-phase and cell division, as well as differences in the regulation of origin licensing and activation events. More recently, it was shown that embryonic stem cells of mice and humans are also subject to massive rearrangements in their DNA replication timing program during development that can affect as much as 50 % of the whole genome.

In the course of this work, I took a closer look at these developmental differences in embryonic stem cells of the mouse (mESCs) and analyzed how they manifest *in situ*. With the use of classical nucleotide pulse-chase experiments, I performed a detailed characterization of the DNA replication program of mES cells, which was subsequently confirmed *in vivo*. I observed remarkable differences in the replication timing of pericentromeric heterochromatin, which may be explained on the basis of differences in chromatin organization of pluripotent cells.

With the use of 3-dimensional *fluorescence in situ hybridization* (3D-FISH), I analyzed the specific replication timing of three major chromosomal tandem repeat elements, i.e. minor and major satellite repeats and telomeres, and identified the Y-chromosome as the last structure to be replicated during S-phase in male cells. I could further show that its duplication occurs in a synchronous manner, similar to that of the inactive X-chromosome of female cells, which suggests a distinct mode of replication that may be specific to these two *inactive* chromosomes.

Using a combination of single molecule and super-resolution microscopy techniques I was able to characterize important molecular parameters of the embryonic stem cell replicon, which allowed me to compare the conservation of this crucial functional unit of DNA replication, with that of somatic cells published in a recent study. These data could indicate further developmental differences in the organization of DNA replication, based on mechanisms that might be conserved between mammalian species, frogs and even flies.

Last but not least, I analyzed the effect of the loss of DNA methylation on DNA replication. While the lack of this important base modification did not interfere with the global progression of the DNA replication machinery, the results show that DNA methylation could be important for the control of DNA helix stability and might have the ability to modulate DNA-dependent metabolic processes on the molecular level.

ZUSAMMENFASSUNG

Die DNA-Replikation ist einer der fundamentalsten Prozesse in jedem lebenden Organismus, der für die Übertragung der (epi)genetischen Information einer Zelle notwendig ist. Um die fehlerfreie Vervielfältigung und Weitergabe des gesamten genetischen Materials genau einmal pro Zellzyklus zu gewährleisten, folgt die DNA-Replikation einem streng kontrollierten räumlich-zeitlichen Programm, welches über viele Arten hinweg identisch ist. Insbesondere in höheren eukaryotischen Organismen muss die Regulation dieses Programms jedoch dynamisch und flexibel genug sein, um die Koordination mit anderen DNA-abhängigen Prozessen wie Transkription und DNA-Reparatur zu ermöglichen. Darüber hinaus zeigten bereits frühe Studien, dass dieses *replication timing program* bemerkenswerte Änderungen während der zellulären Entwicklung durchläuft, welche auf mehreren Ebenen kontrolliert werden können. So findet die DNA Vervielfältigung während der frühen embryonalen Stadien von *Drosophila* und *Xenopus* beispielsweise innerhalb weniger Minuten statt und ist abhängig von der Verfügbarkeit maternaler Faktoren, welche eine schnelle DNA-Synthese in Abwesenheit von Transkription erlauben, sowie einer spezialisierten Organisation des Zellzyklus - gekennzeichnet durch das Fehlen von *gap* Phasen, was den Zellen ermöglicht zwischen S-Phase und Zellteilung zu wechseln - und Unterschieden in der Regulierung des *licensings* und der Aktivierung von Replikationsursprüngen. Vor einiger Zeit wurde zudem gezeigt, dass embryonale Stammzellen von Mäusen und Menschen ebenfalls einer massiven Umorganisation ihres *DNA replication timing programs* während ihrer Entwicklung unterliegen, welche bis zu 50% des gesamten Genoms betreffen können.

Im Zuge dieser Arbeit habe ich diese Entwicklungsunterschiede in embryonalen Stammzellen der Maus (mESCs) genauer untersucht und analysiert, wie sich diese *in situ* manifestieren. Unter Verwendung klassischer *nucleotide pulse-chase* Experimente führte ich eine detaillierte Charakterisierung des DNA-Replikationsprogramms von mES-Zellen durch, welches anschließend *in vivo* bestätigt werden konnte. Ich beobachtete bemerkenswerte Unterschiede im Zeitpunkt der Replikation von perizentromerischem Heterochromatin, was auf der Basis von Unterschieden in der Chromatinorganisation in pluripotenten Zellen erklärt werden könnte.

Mithilfe von 3D *Fluoreszenz-in-situ-Hybridisierung* (3D-FISH) analysierte ich die Zeitpunkte der Replikation dreier wichtiger chromosomaler Tandem-Repeat-Elemente - *minor* und *major satellite repeats* und Telomere. Außerdem identifizierte ich das Y-Chromosom als diejenige Struktur, welche zuletzt während der S-Phase in männlichen Zellen repliziert wird. Ich konnte ferner zeigen, dass seine Duplikation in ähnlicher Weise synchron verläuft wie die des inaktiven X-Chromosoms weiblicher Zellen, was auf eine spezielle Art der Replikation hindeutet, welche für diese zwei *inaktiven* Chromosome spezifisch sein könnte.

Mit einer Kombination von Einzelmolekül- und hochauflösenden Mikroskopietechniken konnte ich wichtige molekulare Parameter des Stammzellreplikons charakterisieren, was es mir ermöglichte, diese wichtige funktionale Einheit der DNA-Replikation mit der von somatischen Zellen zu vergleichen, welche in einer aktuellen Studie publiziert wurden. Diese Daten könnten auf weitere Entwicklungsunterschiede bei der Organisation der DNA-Replikation basierend auf Mechanismen hinweisen, welche zwischen Säugetierarten, Fröschen und sogar Fliegen konserviert sein könnten.

Abschließend habe ich die Auswirkung des Verlusts von DNA-Methylierung auf die DNA-Replikation untersucht. Während das Fehlen dieser wichtigen Basenmodifikation den generellen Verlauf der DNA-Replikationsmaschinerie nicht beeinträchtigte, zeigten die Ergebnisse, dass die DNA-Methylierung eine Rolle in der Kontrolle der DNA-Helix-Stabilität spielen könnte, um so DNA-abhängige Stoffwechselprozesse auf molekularer Ebene zu beeinflussen.

INTRODUCTION

DNA replication - basics & principles

Together with transcription and DNA damage repair, DNA replication is one of the most important metabolic processes in every cell. Incorrect duplication of the genome during DNA replication poses a major threat to every organism and can lead to genomic instability, a potent source of various diseases, including cancer. Thus, each cell within an organism needs to ensure that all its genetic information is properly maintained by guaranteeing that its duplication occurs only once per cell cycle.

Many past and present studies address the cellular and molecular details of DNA replication, leading to our current understanding of how it is organized and regulated in space and time. Nonetheless, many open questions still remain, including most importantly how replication origins are defined in more complex eukaryotic organisms and how exactly epigenetic mechanisms and nuclear chromatin organization are involved in the regulation of different aspects of DNA replication dynamics.

New and improved methodologies have helped to dissected especially the molecular details of DNA replication. However, they not only helped to acquire more knowledge and a better insight but also further uncovered how complex this process really is.

DNA replication can be subdivided into two crucial steps, i.e. **i**) the recognition and subsequent licensing of origins of DNA replication at thousands of sites on each chromosome and **ii**) the ATP-dependent activation of a subset of these sites by cyclin-dependent kinases followed by the duplication of all genetic and epigenetic information by specialized multiprotein complexes.

In the following sections I will provide a summarized overview of the current knowledge about each of these steps.

Origin definition and licensing: formation of the pre-replication complex

Where and how exactly DNA replication initiates along mammalian chromosomes has been of great interest since the earliest experiments and the postulation of the replicon model by Jacob *et al.* (Cairns, 1966; Edenberg and Huberman, 1975). This elegant model, established based on experiments performed with bacteria, proposed that initiation of DNA replication depends on two crucial components, i.e. *cis*-acting elements (*replicators*) that define initiation sites genetically and that are bound by *trans*-acting proteins that act as *initiators* of DNA synthesis (Jacob and Brenner, 1963). Given that a plethora of such *initiator* proteins has been identified throughout all kingdoms of life, this model still holds valid in many respects until today (Masai, 2013; Masai et al., 2010).

In the more than 50 years that have passed since its postulation, DNA replication has been investigated from many perspectives and in the context of other cellular processes, in which the molecular machineries often exhibit overlapping functions and were thus assumed to be involved in the regulation of replication as well. This includes not only primary DNA sequence in prokaryotes and the eukaryot *Saccharomyces cerevisiae* (budding yeast), but also chromatin organization and large scale nuclear architecture as well as epigenetic factors. Not surprisingly, a more or less direct influence of all these factors on DNA replication has been observed in these studies.

In the most basic definition, an origin of replication serves to unwind template DNA at specific sites to allow DNA synthesis to initiate (Burhans et al., 1990). In most viral, bacterial and lower eukaryotic genomes, origins of replication are defined on the DNA sequence level. The first *oris* have originally been discovered in *E.coli* and later studies in the yeast *S. cerevisiae* showed the existence of similar sequences termed *autonomously replicating sequences* (ARS). All ARS elements sequenced to date contain the 11-bp core consensus sequence 5'-(A/T)TTTAT(A/G)TTT(A/T)-3' (*autonomous consensus sequence*, ACS) and when inserted into plasmids, they confer the ability for *in vitro* self-propagation of the DNA fragment carrying the sequence (Palzkill and Newlon, 1988). Budding yeast, however, are the only eukaryotic organism that contains such origin defining sequences. And even here the definition by plain DNA sequence is not enough, as reflected by the finding that binding of reconstituted replication factors *in vitro* is only sequence specific when the template is associated with chromatin proteins and not naked (Devbhandari et al., 2017; Kurat et al., 2017).

With the evolution of the far more complex genomes of mammalian organisms, that are associated with DNA-binding proteins and organized into higher order structures, there was also a need for a different 'origin signature'. This is even more important regarding the fact that metazoan cells need to be able to adapt their choice of replication origins to the highly dynamic environment within the cell nucleus and the dramatic rearrangements that take place in species with different developmental stages (Blow et al., 2001; Hiratani et al., 2008; Newport and Kirschner, 1982; Pope et al., 2010; Ryba et al., 2010). Against this background, even the few well studied mammalian origins to date posses different characteristics.

The *Lamin B2* or *c-Myc* origins, for example, are more localized and site-specific (Abdurashidova et al., 2000; Ghosh et al., 2006), while the human *dihydrofolate reductase* (DHFR) origin carries multiple potential sites within a ~55 kb initiation zone from which DNA replication can initiate (Dijkwel and Hamlin, 1995). Until today, no specific sequence motif has been identified in mammals, that would show properties comparable to that of ARS sequences in yeast and it is unlikely that this will be the case in the future.

Irrespective of the question how mammalian origins of replication are defined, the subsequent licensing steps are getting better and better understood. Interestingly enough, these first steps in the preparation for genome duplication, take place already during late M/G1-phase and, in metazoans, in the absence of a nuclear membrane (Dimitrova and Gilbert, 1999). Several recent studies, involving biochemical experiments with reconstituted replication factors from yeast, recently provided the most detailed insight about the stepwise assembly of the eukaryotic replication fork (Coster and Diffley, 2017; Douglas et al., 2018). In addition to that, cryo-EM analysis of licensed mammalian helicase complexes unveiled an hitherto unexpected organization of the replisome that has strong implications for the exact process of bidirectional DNA replication (Abid Ali et al., 2017; Georgescu et al., 2017; Li and O'Donnell, 2018; Zhou et al., 2017).

Formation of the so called pre-replication complex (pre-RC) starts with the recognition of origins and their binding by the origin recognition complex (ORC) during telophase. The ORC is a hetero-multimeric protein complex consisting of six subunits (ORC1-6), originally identified in budding yeast and for which homologues have been identified in most species, including mammals (Gavin et al., 1995; Gossen et al., 1995; Rowles et al., 1996). Although the mechanism of recognition by the ORC appears to be conserved, the complex shows little or no sequence specificity (Zellner et al., 2007), suggesting that a defined sequence for

DNA replication initiation might not be required at all, supporting the notion of high flexibility of mammalian replication start sites.

In a next step, the complex recruits the AAA^{+}-ATPase CDC6 (*cell division control protein 6*), which is thought to modulate the affinity of ORC for certain subsets of sequences that can generally serve as DNA replication origins (Harvey and Newport, 2003; Mizushima et al., 2000; Speck and Stillman, 2007). CDT1 bound to the putative mammalian helicase *minichromosome maintenance complex* (MCM2-7) then gets recruited to the ORC/CDC6 complex, completing pre-RC formation. The ORC/CDC6 complex probably serves as a clamp loader that opens up MCM2-7 and closes it around the double-stranded DNA (dsDNA) template, as has been shown in (Shin et al., 2003). However, the exact molecular mechanism and also the number of pre-assembled ORC complexes is currently under debate (Ticau et al., 2015).

The two MCM complexes are loaded as inactive double hexameres in a head-to-head orientation and ATP hydrolysis by MCM releases CDT1 from the complex (Coster et al., 2014; Kang et al., 2014). After these steps ORC, CDC6 and CDT1 all become dispensable for subsequent origin firing when the cell enters S-phase (Arias and Walter, 2005; Shibata et al., 2016).

Licensing of replication start sites occurs in excess, a phenomenon referred to as the *MCM paradox* (Hyrien et al., 2003) and this '*overloading'* of DNA with helicase complexes serves as a safety mechanism for the cell. Dormant origins can become activated in the case of fork stalling events in which individual replication forks emanating from activated origins experience some sort of replicative stress that leads to replication pausing (Ge et al., 2007). All steps defining pre-ORC formation are summarized in Fig. 1 (A-D)

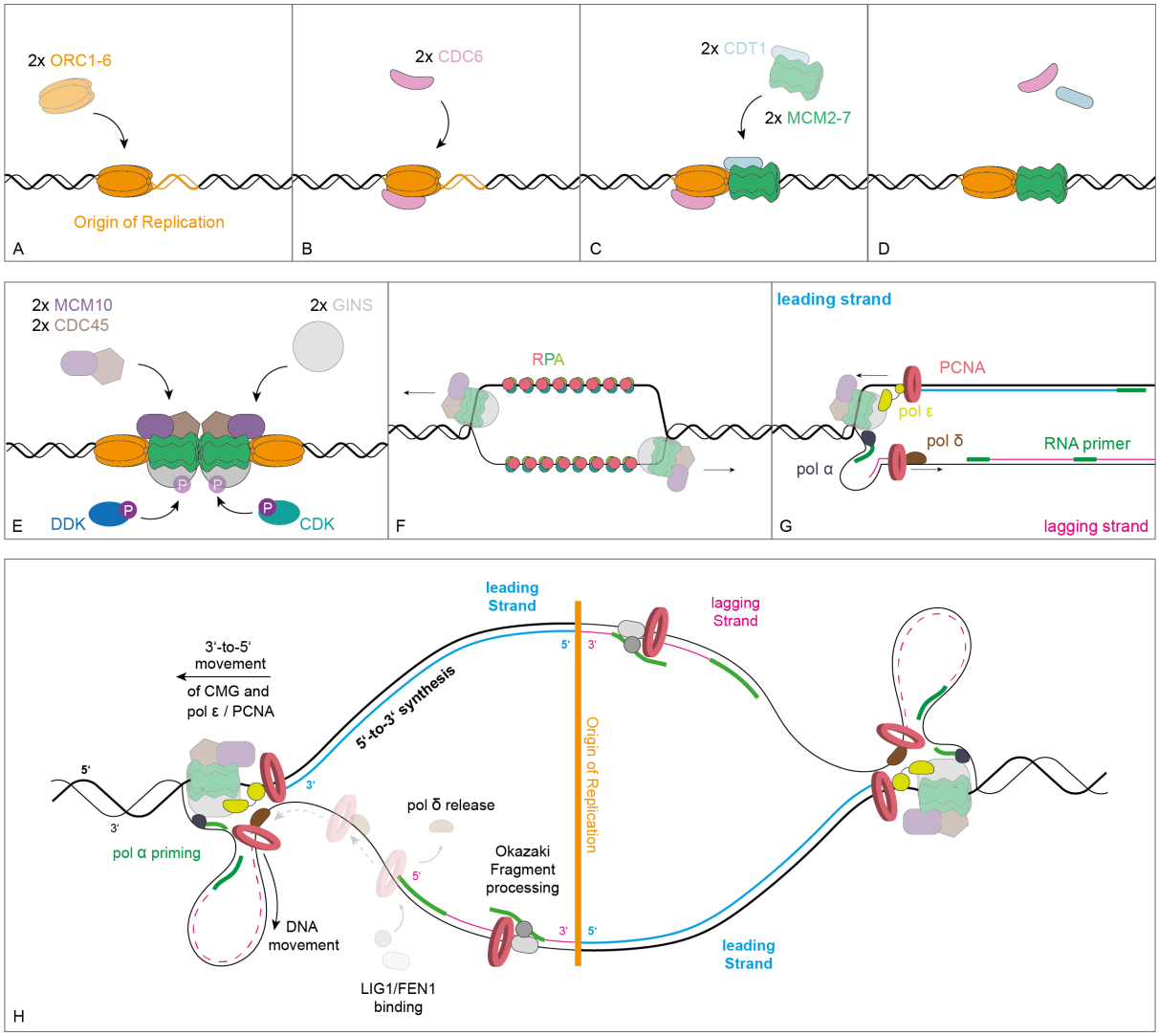


Figure 1 - From origins to replication forks: an overview of the individual steps from origin licensing to replication complex activation upon S-phase entry. In the first step, two ORC complexes (ORC1-6) assemble to potential origin of replication (**A**). In mammals, were origins are not defined by DNA sequence, ORC binding occurs flexibly. ORC then recruits the AAA^{+}-ATPase CDC6 (**B**) which is thought to modulate the affinity of ORC for certain DNA sequences that serve as potential origins. Together they recruit CDT1 which is bound to the MCM2-7 helicase complex, completing pre-RC formation with loading of the two inactive MCM complexes in a head-to-head orientation (**C**). ATP hydrolysis by MCM releases CDC6 and CDT1 (**D**). All factors, including ORC, become dispensable for subsequent origin activation. CDK- and DDK-dependent phosphorylation of MCM subunits lead to the recruitment and loading of CDC45, which stimulates the helicase activity of MCM2-7, and GINS whos exact function is yet unclear (**E**). Together they build the CMG helicase which initiates double-helix unwinding upon phosphorylation. Imposed by the fact that they are bound in a head-to-head orientation, the CMG complexes have to shift from double- to single stranded DNA (ssDNA) in order to pass each other in 3'-to-5' direction (**F**). The single strand binding protein RPA binds to and stabilizes resulting single-stranded regions, preventing the reformation of double-strands or secondary structure and promotes further unwinding. For nascent strand synthesis polymerases alpha (pol α), delta (pol δ) and epsilon (pol ε) are recruited to the replication bubble. Pol α has *de novo* primase activity and synthesizes short RNA primers on the leading (blue) and lagging strand (magenta). From these primer sites, pol δ (lagging strand) and pol ε (leading strand) perform the bulk DNA synthesis in 5'-to-3' direction. The processivity of both polymerases is increased

by the ring-like homotrimeric protein PCNA (**G**). (**H**) is a representation of the organization of all involved replication factors within a replication bubble. The CMG helicase and associated replisome components move in 3'-to-5' direction beginning at the origin of replication in the middle of the bubble structure (orange line). This results in 5'-to-3' DNA synthesis (see left side). As this directionality cannot be achieved for both template strands at the same time, the lagging strand is believed to be looped out from the replisome. Pol δ synthesizes new DNA until it encounters a RNA/DNA primer hybrid site, marking the beginning of a new Okazaki Fragment. PCNA serves as a loading platform for additional proteins involved in DNA replication including *Ligase 1* (LIG1) and *Flap endonuclease 1* (FEN1) for Okazaki Fragment maturation or *DNA methyltransferase 1* (DNMT1, not shown) that maintains DNA methylation patterns in a replication dependent manner.

Activation of licensed replication origins

The successful establishment of pre-replication complexes throughout the genome is followed by their transition to active replication complexes (RCs) upon S-phase entry. The addition of activating phospho-groups to several phosphorylation motifs in MCM subunits by *cyclin-dependent kinase* (CDK) as well as Dbf4-dependent kinase *CDC7* (also named DDK) leads to the stepwise recruitment and stable binding of additional factors to the MCM hexamers. Phosphorylation of the N-termini of the MCM subunits MCM4/MCM6 by DDK is responsible for the recruitment of CDC45 (Fig. 1E) which stimulates the helicase activity of the MCM complex and is additionally also required for DNA helix unwinding (Deegan et al., 2016).

At the onset of S-phase the GINS complex is loaded onto replication origins which depends on CDC7 and CDK (Fig. 1E). GINS stands for 'Go-Ichi-Ni-San' (5-1-2-3) owing to its four subunits Sld5, Psf1, Psf2 and Psf3. Like many other essential replication factors, the GINS complex has initially been identified in yeast, but homologues have also been identified in *Xenopus*, mice and even humans. The exact function of the complex is yet unclear but in *Xenopus* and yeast, it has been shown to travel with the replisome when origins become active and it might be responsible for the recruitment of additional factors to the replication fork, including replication protein A (RPA, (Kanemaki and Labib, 2006; Labib and Gambus, 2007)).

Altogether, these factors form the eleven-subunit CMG complex (CDC45-MCM2-7-GINS) which represents the actual eukaryotic helicase complex. In this conformation, the complex is still bound to double-stranded DNA (dsDNA). Upon activation by CDK and DDK (Fig. 1E), a mechanism that is still poorly understood, the CMG complex promotes the initial unwinding of the replication bubble and shifts from double-stranded to single-stranded DNA (ssDNA) (Fig. 1F; (Douglas et al., 2018)). The heterotrimeric protein complex RPA, which is

the major single-strand binding (SSB) protein in eukaryotes and consists of the three subunits RPA1, RPA2 and RPA3, stabilizes the single-stranded regions (Fig. 1F). This promotes further unwinding and prevents renaturation of the individual strands back into the double-helix structure as well as the formation of secondary structures that could block replication fork passage (Walter and Newport, 2000).

The actual synthesizing machineries of the replisome, the polymerases alpha (pol α), delta (pol δ) and epsilon (pol ε) are recruited as a results of the binding of CDC45 and RPA. DNA polymerase α itself consists of four subunits with differing catalytic activity and is the only one of the three polymerases capable of starting DNA synthesis *de novo*. Thus, pol α synthesizes the short RNA/DNA hybrid primers required for both leading- and lagging-strand synthesis, which are subsequently extended up to 30 nucleotides (Fig. 1G + H).

Both polymerase δ and polymerase ε synthesize DNA in 5'-to-3' direction and exhibit proof-reading activity as well as higher processivity compared to polymerase alpha. With these properties, the enzymes are responsible for leading- (pol ε) and lagging-strand synthesis (pol δ), respectively (Fig. 1G + H; (Miyabe et al., 2011)). The increased processivity, however, is coupled to their association with *proliferating cell nuclear antigen* (PCNA; (Fig. 1G + H); (Chilkova et al., 2007)), a homotrimeric ring protein loaded onto chromatin by the clamp loader *replication factor C* (RFC). PCNA serves as a loading platform for all sorts of proteins required for DNA replication and is also involved in other molecular processes such as DNA repair and translesion synthesis (Boehm et al., 2016; Sporbert et al., 2005).

In theory, the total number and spacing of activated origins as well as the speed of the two replication forks travelling bidirectionally from these sites, are sufficient parameters to define the total time required to fully replicate a genome. Practically, however, the situation is much more complex in the context of the highly dynamic nuclear environment in mammals. Consequently, the exact dynamics of the many factors involved at the replication fork are extraordinarily difficult to study and, thus, are not yet fully understood.

One particular problem is posed by the fact that both, the leading and lagging strand, need to be synthesized simultaneously by replication forks emanating from each origin. The overall architecture of the eukaryotic replication machinery could allow us to address some of the most important questions that arise in the light of this problematics. Recently published studies dissecting exactly this molecular architecture, provided data that lead to a better, and in parts, new understanding of the dynamics of leading and lagging strand

synthesis (Coster and Diffley, 2017; Douglas et al., 2018; Georgescu et al., 2017; Li and O'Donnell, 2018).

Once the DNA helix has been unwound at the origin, the two helicase complexes begin to move away from each other. It was shown that the subunits of the MCM complex are organized in a way such that both the N- and C-terminal domains of all subunits form distinct tiers (N-tier and C-tier, respectively) (Enemark and Joshua-Tor, 2008; Thomsen and Berger, 2009). The C-terminal domains contain ATP binding sites and represent the actual motor of the helicase complex that promotes DNA unwinding and helicase translocation. In eukaryotes, CMG translocates in 3'-to-5' direction and it is accepted that the separation of both strand by the helicase occurs via a strand-exclusion mechanism, where one strand is encircled by the complex while the other becomes excluded from its core by sterical hindrance (Singleton et al., 2007).

PCNA and the other replication factors, like the DNA-dependent polymerases, are then either assembled onto each template strand or interact directly with the CMG complex. Cryo-EM single-particle data of the CMG helicase of *S. cerevisiae* bound to DNA published by *Georgescu et al.* provided structural information, and revealed that the polymerase alpha/primase complex and polymerase epsilon are associated on opposite sides of the GINS complex (Fig. 2) (Georgescu et al., 2017). This conformation suggests, that the lagging strand that gets excluded from the helicase core which would be readily accessible for priming by pol α/primase. Pol ε consists of to separated domains, aC-terminal domain (CTD) and a catalytically active N-terminal domain, that are presumably connected by a flexible linker region. This would allow immediate capturing of the leading strand treading out of the CMG helicase for continuous synthesis. In support of this organization of the two strands with the helicase, *Fu et al.* showed that the helicase is capable of bypassing blocks placed on the lagging, but not the leading strand (Fu et al., 2011).

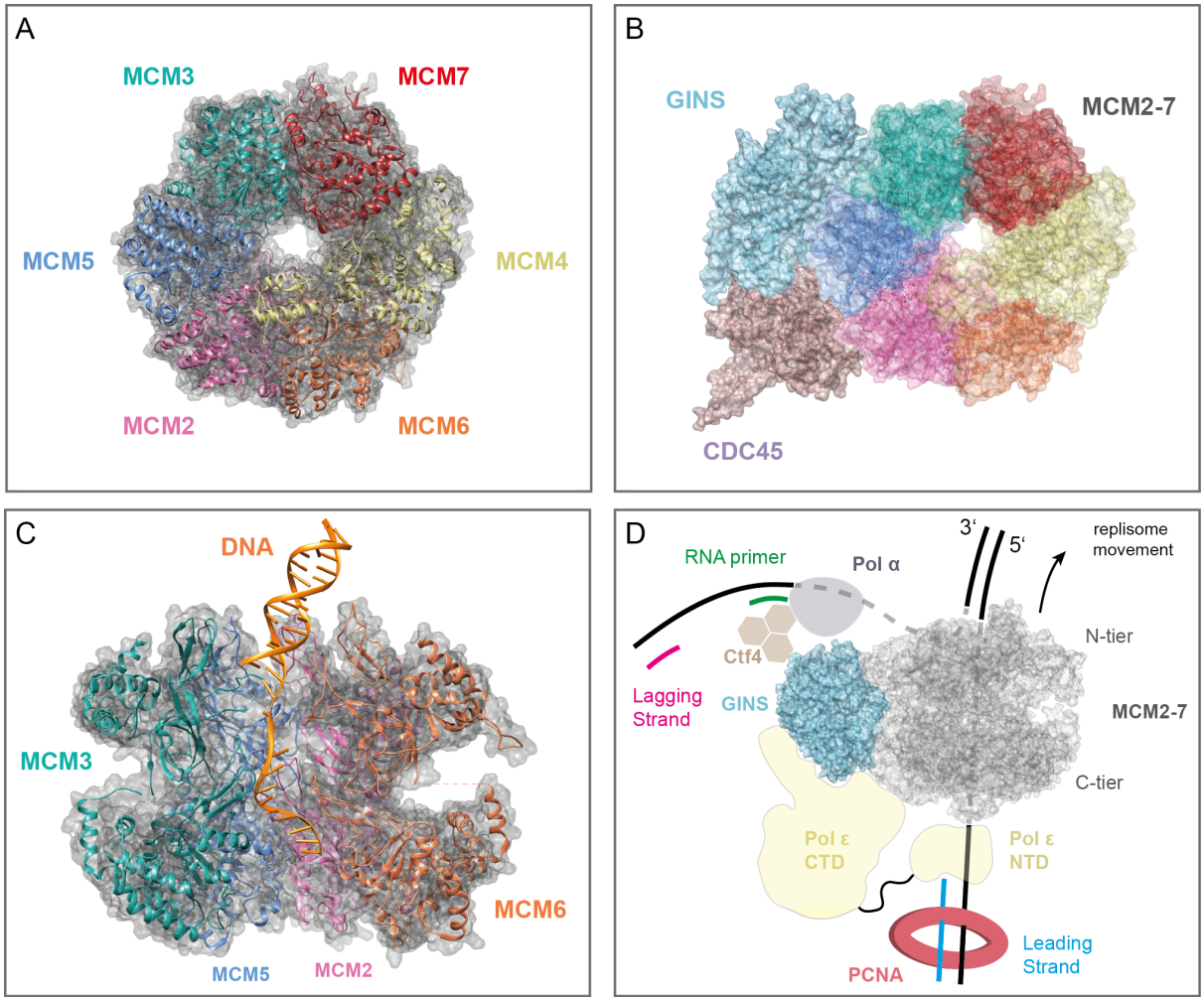


Figure 2 - Knowledge about the molecular architecture of the eukaryotic replisome provides important mechanistic insights. (**A**) Crystal structure of the eukaryotic MCM2-7 helicase complex from *S. cerevisiae* (PDB: 5U8S, (Georgescu et al., 2017)). Shown are the six individual subunits that together form the ring-shaped helicase complex. For full helicase activity, MCM2-7 further associates with CDC45 and GINS forming the so-called CMG helicase complex (**B**). (**C**) Top view of the CMG complex (MCM subunits 4 and 7 removed), depicting how unwound, single-stranded DNA (orange) enters the central channel of the complex upon its activation. (**D**) The synthesizing machinery assembles onto the CMG complex during S-phase for leading- and lagging strand synthesis. The crystal structure data published by *Georgescu et al.* suggests that MCM2-7 is oriented with its N-tier ahead. Movement of the helicase along the DNA (black line(s)) is maintained through the motor activity of the AAA+ C-tier, pushing the whole complex forward in 3' to 5' direction. The lagging strand is unwound from the leading strand in front of the N-tier and becomes located immediately adjacent to polymerase α, which is associated with GINS through the trimeric protein Ctf4. This conformation allows continuous priming (green) of the lagging strand (magenta) by the polα/primase complex. Polymerase ε, schematically represented in yellow, is located on the lower part of the C-tier of MCM, bound via its C-terminal domain (CTD). The N-terminal domain (NTD), thought to be connected to the CTD via a flexible linker, bears the catalytic activity required for DNA synthesis. When the leading strand exits the central channel of MCM2-7 it is captured by the catalytic subunit of polymerase α, available for immediate synthesis. PCNA travels behind the complex, increasing the processivity of pol ε.

Another long standing question that has remained unanswered, largely due to the lack of sufficient methods, is how many PCNA molecules are assembled onto both the leading and lagging strand. Also, where exactly along the DNA loading of PCNA occurs, how long it remains associated with the DNA and how potential loading of new PCNA and unloading of existing trimers by RFC would be achieved during replication, are related questions. Using enrichment of protein-associated nascent DNA followed by sequencing (a method termed *eSPAN*), *Yu et al.* were recently able to show that the polymerases epsilon and delta are distinctly enriched on the leading and lagging strand, respectively, reflecting their suggested role in the synthesis of either of the two strands (Yu et al., 2014). Furthermore, the authors showed that PCNA is also specifically enriched on the lagging strand during DNA replication in budding yeast. This suggests release of the PCNA trimer at every Okazaki fragment primer site and constant reloading of new PCNA from the nucleoplasmic pool to proceed with the helicase. In higher mammalian species, however, PCNA turnover at the replication forks seems to be very limited as shown by FRAP studies performed by *Sporbert et al. (Sporbert et al., 2005).* The authors could show that the processivity factor was exchanged only up to maximally 30%, indicating that it remained associated with the DNA most of the time and that release from the replication fork for Okazaki fragment maturation and loading of new PCNA to the strand, was unlikely.

The *E.coli* β clamp was proposed to slide along DNA during DNA replication(De March et al., 2017) and although PCNA has strong structural similarity with the bacterial clamp complex, a direct test of such sliding mechanism for PCNA has not been shown. Thus, another idea that has remained elusive is how, or whether at all, the whole replication machinery or components thereof slide along the template strands during S-phase.

Otherwise, the replicated DNA might get reeled through sister replisomes (the two replisomes of a bidirectional fork emanating from the same origin), that remain in a more or less fixed position. In support of this, one study in yeast showed that sister replisomes remained in close proximity during DNA replication, while the duplicated DNA got spooled through the complexed machinery, a finding that has already been observed in prokaryotes ((Kitamura et al., 2006), and references therein).

Together with past experiments, this allowed us to formulate concepts that describe how the cellular machineries can ensure duplication of the entire genome once per cell cycle and in the context of the complex chromatin architecture and time restrictions in mammalian cells.

DNA replication in situ

The initial licensing of replication origins in excess is followed by the sequential activation of only a subset of these start sites over the 8-10 hours of S-phase in mammalian cells. Many of the features of DNA replication show a high similarity in different species including yeast, flies, mice and humans (Jackson and Pombo, 1998; Leonhardt et al., 2000; Ma et al., 1998; Manders et al., 1992; O'Keefe et al., 1992), and homologues for the key factors of replication described above, have been identified in most of these species (Fragkos et al., 2015). This conservation and the fact that the individual steps are carried out in a non-random manner, manifests in the existence of a conserved replication timing program. Generally, this program correlates well with transcriptional activity, where euchromatic gene-rich sequences replicate early and AT-rich, gene-poor heterochromatic segments replicate late in S-phase (Woodfine et al., 2004).

The application and detection of nucleotide analogues or replisome components in fixed cells is a useful method to visualize sites of newly synthesized DNA *in situ*, usually referred to as *replication foci* (RFi). Such experiments reflect the above mentioned conservation and uncovered the existence of distinct *replication patterns* (Chagin et al., 2010; Flockinger et al., 1967; Jackson and Pombo, 1998; Leonhardt et al., 2000; Manders et al., 1996; van Dierendonck et al., 1989). These patterns are represented by the dynamic intra-nuclear distribution of the replication signals over the course of S-phase. In metazoan cells three such patterns are frequently observed, and are termed *early* (I), *mid* (II) and *late* (III) S-phase, that correspond to the replication of euchromatin (or R-bands), as well as, facultative and constitutive heterochromatin, respectively (Fig. 3; (Nakamura et al., 1986; Nakayasu and Berezney, 1989; O'Keefe et al., 1992)).

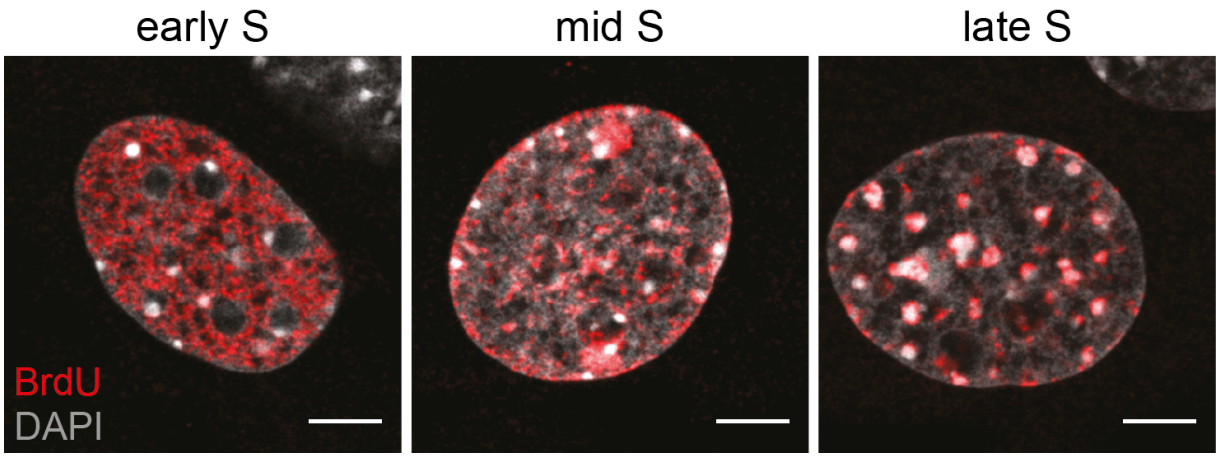


Figure 3 - Replication patterns in somatic cells represent spatio-temporal DNA replication dynamics. Typical replication patterns seen in replicating somatic cells after labeling and immunodetection of nucleotide analogues (e.g. BrdU, red) into nascent DNA. Euchromatin is replicated at the beginning of S-phase (early S) and marked by homogeneous distribution of replication signals throughout the nucleus. Pericentromeric heterochromatin (bright spots) and the nuclear and nucleolar periphery are unlabeled at this early stage. Facultative heterochromatin replicates in the middle of S-phase (mid S), marked by perinuclear and perinucleolar distribution of replication signals. In the cell type shown here (female mouse myoblast cells) two additional structures are marked by strong accumulation of replication signals that represent the inactive X-chromosomes (two in this case, as the cell line is tetraploid). In late S-phase, constitutive pericentromeric heterochromatin is replicated, which forms clustered structures in mouse cells called *chromocenters*. Scale bar = 5 µm

Interestingly, early pulse-chase experiments showed that the distribution of the labeled nascent DNA within the cell nucleus appears to be stable even over multiple cell cycles. This suggested the existence of a conserved chromatin organization that could be important for the regulation of cellular processes, including DNA replication (Jackson and Pombo, 1998). Indeed, it was shown that the position of replication foci is already established at the repositioning and anchoring events that take place during the so called *timing decision point* during early G1-phase (Dimitrova and Gilbert, 1999; Sadoni et al., 2004).

To build up on these findings, genome-wide techniques have been developed and applied to study the DNA replication timing program in different species and cells of varying developmental origin (Hiratani et al., 2008; Ryba et al., 2010). Those studies confirmed that the coordinated activation of clusters of replication origins, as has been proposed from single-molecule experiments (Berezney et al., 2000), occurs in large chromosomal segments. These become visible as multimegabase domains and were consequently termed *replication domains* (RDs). How exactly these domains relate to replication foci and

how they fit to the concept of an underlying organizational unit, as described in the following section, is not yet clear.

From these and many other pionieering studies, two major models evolved that aim to describe the organisation of replication machineries within the cell nucleus and how they interact with chromatin during S-phase. The first model is based on electron microscopic studies of embedded HeLa cells and proposes that replisomes become attached to a diffuse underlying *nuclear matrix* or *scaffold* during S-phase (Hozak et al., 1994). The existence of such scaffold, however, is still debated. In this model, replication machineries represent fixed granular entities, termed *replication factories,* that become visible as replication foci, e.g. when replisome components are detected with specific antibodies. To achieve proper duplication of the DNA at those immobile sites, DNA would have to be reeled through the factories, with each of them processing clusters of replicons that are activated together at distinct times during S-phase. Upon exits of the DNA from a factory, it becomes organized into distinct chromatin loop structures.

More recent FRAP experiments, however, have shown that replisomes assemble *de novo* adjacent to previously activated sites (Sporbert et al., 2002), suggesting that replisomes do not persist as permanent factories throughout an entire S-phase. Instead, such *next in-line* activation represents the basis for the second model, in which DNA replication is carried out in a ‘domino-like’ fashion. Hereby, the completion of a full round of DNA replication relies on the combination of **i)** the initial stochastic activation of a few origins at random sites along the DNA with high firing probability and **ii)** the subsequent domino-like activation of adjacent sites with decreasing firing efficiency at later time-points during S-phase (Friedman et al., 1997; Lob et al., 2016; Yamashita et al., 1997). The propagation of activating signals may hereby be achieved by localized changes in chromatin compaction, maybe as a results of the replication process itself. Although this model has been supported by studies that showed that the continuous activation of new replication sites correlates well with their linear organization along chromosomes (Braunstein et al., 1982; Maya-Mendoza et al., 2010), there is no direct experimental evidence for an exact mechanism as of today.

Chromatin organization in the context of DNA replication

Accomplishing duplication of all genetic material is particularly challenging in higher organisms like mammals, in which the individual multi-megabase DNA fibers are organized at multiple levels into a set of chromosomes that distribute non-randomly in the interphase nucleus. Chromosome organization begins with the association of DNA with the core histone octamer consisting of two H2A-H2B and H3-H4 histone dimers (Hubner et al., 2013). Each histone core particle is wrapped with 147 bp of DNA, while another 10 - 80 base pairs connecting the individual octamers are bound by a special histone variant, the linker histone H1.

In this complex, the chromatin fibers has a diameter of ~10 nm and is often referred to as *beads-on-a-string.* The next level of compaction is achieved by the interaction of the individual `histone beads` with each other that, resulting in a 30 nm-fiber, a structure that has been observed multiple times *in vitro* but seldomly *in vivo* and which existence is debated until today. Due to the flexible interactions between the histones, chromatin in the interphase nucleus is believed to exist in a dynamically disordered and interconnected state (Maeshima et al., 2010). In that, transient interactions between the individual chromosomes result in packaging as distinct and more or less separated chromosome territories (Cremer and Cremer, 2001; Cremer and Cremer, 2010). Only with the generation of such highly compacted structures, it is possible to achieve sufficient storage within the very limited volume of a cell nucleus at a ratio of ~1000-fold for interphase chromosomes and an even higher degree for mitotic chromosomes (~factor 10; (Alberts, 2008)).

Replicons, i.e. the segment of DNA replicated a pair of bidirectional replication forks from a single origin of replication, are organized and activated in clusters, as demonstrated in early autoradiographic fiber studies of single DNA molecules (Jackson and Pombo, 1998). These studies suggested that individual clusters are on average 1 megabase (Mb) in size and consist of 2 - 9 smaller replicons of, on average, 100 - 200 kb (Berezney et al., 2000; Huberman and Riggs, 1968; Jackson and Pombo, 1998). Studies that used the 'DNA halo technique', reported that the sizes of replicons are in good agreement with measured sizes of chromatin loops, indicating that loop structures represent the DNA element that defines replicons as the functional unit in the context of DNA replication (Buongiorno-Nardelli et al., 1982; Marilley and Buongiorno-Nardelli, 1984).

Nakamura et al., performed the first experiments to show that when cells were labeled for increasing length with nucleotide analogues like BrdU, these replicons become visible as

the replication foci described earlier (Nakamura et al., 1986). Based on the brightness of the signals obtained in their initial labeling experiments, the authors calculated a DNA content of at least 1 Mb per focus. Thus, it was concluded that the replication foci observed in interphase nuclei represent the replicon clusters seen in DNA fiber experiments.

Importantly, depending on the level of optical resolution (wide-field, confocal or super-resolution) highly variable numbers of replication foci have been reported in different studies, resulting in a high variability of up to 100-fold (Reinhart and Cardoso, 2017). Also, the measurement of replicon sizes usually results in a broad distribution of 30-450 kilobase pairs (Berezney et al., 2000). This can be explained, in part, by technical limitations, e.g. due the inherent resolution limit of DNA (autoradiographic) fiber based experiments, in which very small or large replicons might not be measured accurately enough (Techer et al., 2013).

The focal signals that results after labeling of nascent DNA are of varying sizes and their spatial position is constantly changing over the course of S-phase. However, the general foci size distributions do not significantly change, even when the labeling time is significantly increased (Ma et al., 1998). These observations, in combination with the observed persistence of the spatial distribution of replication foci over several cell cycles, suggested an important connection to the underlying chromatin organization and the interest in addressing this relationship grew rapidly. Many recent studies, although mostly outside the context of DNA replication, began to dissect how chromatin loops may be established and maintained within the interphase nucleus and how chromatin (loop) organization is involved in the regulation of DNA-dependent processes.

The loop structure itself is a well known element of chromatin (van Driel et al., 2003) and provides well suited features for both the stability of the DNA fiber as well as the possibility for long range interaction between distantly located sequences. The latter forms the actual basis for the regulation of DNA metabolic processes, e.g. as seen in the interactions between the enhancer and promoter of the β-globin gene, which are ~50 kb apart (Wijgerde et al., 1995).

Suitable approaches to study and characterize the frequent loop formations within cells are the measurement of the short and long range chromosomal interactions by *chromosome conformation capture* methods (Lando et al., 2018; Stevens et al., 2017). Such genome-wide mapping of chromosomal interactions, revealed that the chromosomes of higher eukaryotes are functionally compartmentalized into distinct *topologically associating*

domains (TADs, (Dixon et al., 2012)). The sequences within these domains are represented by frequently self-interacting regions that are visualized in so called Hi-C contact maps (Dekker and Mirny, 2016; Dixon et al., 2012). The computational modelling of interphase chromosome organization aims to describe these interactions mechanistically. Models of loop extrusion mechanisms, in which the DNA is extruded through a connected pair of *loop extrusion factors* (LEFs) to form large but dynamic loops, sufficiently recapitulated the contact maps for TADs obtained in Hi-C studies (Fudenberg et al., 2016). Good candidate factors that may be responsible for *in vivo* loop formation, are cohesins and the insulator protein CTCF, as they exhibit all molecular functions required to carry out this task.

Cohesin is a ring-structured multiprotein complex consisting of the subunits Smc1 and Smc3, both members of the *structural maintenance of chromosomes* protein family. Cohesin has originally been identified and studied in the context of sister chromatid cohesion during the transition from S- to M-phase of the cell cycle (Guacci et al., 1997; Michaelis et al., 1997; Nasmyth and Haering, 2009), but was also already shown to participate in DNA repair (Sjogren and Strom, 2010) and the regulation of gene expression (Hadjur et al., 2009; Hou et al., 2010; Nativio et al., 2009). Loading of cohesin onto DNA occurs at the M/G1 transition and depends on a heterodimeric complex of the proteins Scc2 and Scc4 and, after that, it is able to translocate along DNA to different genomic sites. Furthermore, it has structural similarities to known motor proteins (see references in (Fudenberg et al., 2016)).

CTCF is an ubiquitously expressed transcription factor (Filippova et al., 1996) and well characterized insulator protein. Sequencing studies revealed that several thousands of CTCF binding sites of the consensus sequence 5'-CCGCGNGGNGGCAG-3' are present throughout the genome. Based on its insulator function, CTCF was one of the first factors shown to promote chromatin loop formation between its oppositely oriented binding sites (Handoko et al., 2011; Splinter et al., 2006). Indeed, several studies could already proofed that both factors, CTCF and cohesin, bind to the same sites on DNA (Parelho et al., 2008; Wendt et al., 2008) and to promote loop formation (Hadjur et al., 2009; Hou et al., 2010; Nativio et al., 2009).

A study published by Guillou and colleagues further investigated the role of cohesin during DNA replication and could show that it is enriched at DNA replication origins and even physically interacts with the MCM2-7 complex (Guillou et al., 2010). Knock-down (KD) of cohesin resulted slower S-phase progression.The authors also compared DNA halo sizes, which were significantly increased after cohesin KD and was already shown to affect S-

phase progression in a previous study (Courbet et al., 2008). Interestingly, cohesin knockdown also lead to a decrease in the number of activated origins, reflected by an increased inter-fork distance seen in DNA fiber experiments. Based on these results, the authors concluded that cohesin is important for the formation and/or stabilization of chromatin loops that bring together clusters of replicons during DNA replication. In the absence of cohesin, the connection at the basis of the replicon-containing loops is disrupted and leads to a structural reorganization, resulting in fewer replicons on larger loops that ultimately lead to fewer initiation events and, therefore, larger replicons.

Together, these results propose a loop-based unit of chromatin organization, potentially mediated by cohesin or functionally related proteins, and that these units become visible in the form of replicons in the context of DNA replication.

DNA replication during embryonic development

Beyond the effort to dissect the different steps of DNA replication and the architecture of the involved molecular machineries, there is great interest in addressing the differences and changes in the spatial or temporal order of the DNA replication program that occur during development.

The best studied examples are the early developmental stages of *Drosophila* and *Xenopus* embryos in which cell division occurs within 5-20 minutes, respectively (Blumenthal et al., 1974; Hyrien et al., 1995; Hyrien and Mechali, 1993), while the same process takes several hours in differentiated cells (Nordman and Orr-Weaver, 2012). This rapid S-phase behaviour is accompanied by a substantially different spacing and activation of replication origins and with that differences in the resulting replicon sizes (Hyrien et al., 1995; Hyrien and Mechali, 1993; Walter and Newport, 1997). Generally, such massive rearrangements emphasize the need for flexibility as reaction to a dynamically developing environment. In *Xenopus* and *Drosophila* this flexibility has been linked to the absence of transcriptional activity, for which the organisms compensate by the deposition of maternally derived stockpiles of all sorts of factors required for development (Edgar and Schubiger, 1986; Newport and Kirschner, 1982), and only upon differentiation and the re-establishment of gene transcription, do the sites of replication initiation become more restricted to intergenic spacers (Maric et al., 2003; Sasaki et al., 1999).

The rapid cell division that are a characteristic feature of many metazoan species could actually represent a necessary adaptation of their freely developing eggs in order to cope with the different environmental threats to which they are exposed. Mammalian organisms, on the other hand, are less affected by such exogenous factors, hence they do not exhibit such fast cell division rates, but do otherwise show some interesting overlap in the organization of their early embryonic cell cycles. The latter include, for example, short gap phases, at least during some developmental stages and the absence of transcription in the first zygotic cleavage stages of mice (Fragkos et al., 2015; Sansam et al., 2015).

Although most of our current knowledge comes from studies in frogs and flies, it has been shown that specific replication patterns exist as early as the 1-cell stage in mouse embryos (Ferreira and Carmo-Fonseca, 1997) and that the replication timing program of a developing cell undergoes large rearrangements during lineage commitment. These changes affects as much as 40% of the entire genome (Hiratani et al., 2008; Rivera-Mulia et al., 2015). A similar reorganization of replication domains has been found in several independent studies that analyzed both human and mouse embryonic stem cells (Hiratani et al., 2008). Future studies will likely reveal further interesting differences and similarities of the DNA replication program between different species

Regulation of DNA replication by epigenetic mechanisms

Combinations of different epigenetic modifications define distinct *chromatin signatures*, that have been linked to the manifestation of replication timing programs of cells. Besides post-translational modifications of histone tails, the exchange of canonical histones with specialized variants (Sansoni et al., 2014), different states of DNA methylation, the binding of chromatin-proteins and -modifiers as well as small nuclear RNA molecules (Aladjem, 2007), have entered the stage of DNA replication. The isolated study of their exact function in the regulation of replication, however, has proven difficult to dissect because of the strong crosstalk between these epigenetic mechanisms and the different processes that they regulate.

Given that chromatin structure is a very prominent candidate for the regulation of DNA replication timing, the general believe is that epigenetic modifications play a central role in modifying chromatin structure globally or locally, thereby influencing the accessibility for factors required for the licensing and activation of replication origins (Aladjem, 2007). This makes sense especially in the context of mammalian genomes, on which origins of

replication are not defined by a particular sequence and where the choice of replication start sites occurs anew at every G1-phase, followed by the activation of only a subset of all licensed sites.

DNA methylation describes the covalent addition of methyl groups to the C5 position of cytosines, which then becomes 5-methylcytosine (5mC) as originally described in (Hotchkiss, 1948). Addition of the methyl-group preferentially occurs at cytosines of CpG dinucleotides, and arrays of these dinucleotide motif, so called CpG islands (CGI, ~1000 bp on average) are found at promoter regions of most genes in mammalian organisms (~70%, (Saxonov et al., 2006)). Interestingly, however, promoter CGI are usually unmethylated and transcriptional regulation is achieved by methylation of CpGs located directly within the genes. Besides its implications in gene silencing, DNA methylation is also important for the inactivation of one of the female X-chromosomes as well as transposable elements and plays a role in other physiological processes, including embryogenesis and imprinting of genes during development. Consequently, alterations in DNA methylation can lead to the misregulation of gene expression and development of cancer and related diseases (Gopalakrishnan et al., 2008; Jones and Baylin, 2002; Li, 2002; Robertson, 2005).

Cytosine methylation is established by specific DNA methyltransferases (DNMTs) that use S-adenosyl methionine as a methyl-donor. In mice, three different DNMTs exist. DNMT3A and DNMT3B, respectively, are responsible for *de novo* establishment of DNA methylation during embryonic development (Bestor et al., 1988; Li et al., 1992). DNMT1, on the other hand, has a preference for hemi-methylated DNA and maintains the methylation patterns in a DNA replication dependent manner (Leonhardt et al., 1992; Song et al., 2011). Studies showed that DNMT1 can be targeted to DNA replication foci by two different domains, which are the PCNA binding domain (PBD; (Leonhardt et al., 1992)) and the polybromo-1 protein homologous domain (PBHD; (Liu et al., 1998)). In addition, recruitment to post-replicative pericentromeric heterochromatin is achieved by a different mechanism (Easwaran et al., 2004).

Besides its above mentioned roles, DNA methylation appears to also have direct regulatory function in the context of DNA replication. The highly methylated constitutive heterochromatic regions, imprinted genes and the inactive homologue of the female X-chromosome, which contains highly methylated promoter regions, all appear to be replicated preferentially late in S-phase. However, this observation is by no means without exception and changing the methylation status of imprinted genes by treatment with the demethylating agent 5-azacytidine (5-aza) did not result in an measureable switch in their

replication timing (Bickmore and Carothers, 1995). Interestingly, decreased methylation seen in teratocarcinoma cells from mice, led to earlier replication of their heterochromatic compartment (Selig et al., 1988). This is in line with another study, in which the loss of DNA methylation at pericentromeric heterochromatin in embryonic stem cells upon deletion of the DNA methyltransferase DNMT1 or prolonged treatment with 5-aza, also resulted in an advanced replication timing (Jorgensen et al., 2007). In addition, the same study assessed replication timing changes of multiple genetic loci in different knockout mES cell lines lacking various DNA or chromatin modifiers. Their results showed that changes in replication timing were not particularly linked to loss of histone or DNA methylation, but correlated well with changes in histone acetylation levels.

Indeed, more direct modulations of chromatin accessibility occurs at the nucleosomal level either by the post translational modification of histone tails, the incorporation of histone variants, nucleosome remodelling by specialized remodelling complexes or the recruitment of specialized chromatin binding factors (Felsenfeld and Groudine, 2003).

The reversible addition of acetyl-groups to lysine residues of histone tails is thought to alter the degree of chromatin compaction by neutralizing the positive charge of the histone tail, thereby lowering the interaction between core histones. Acetylation of specific sites as well as global histone acetylation levels have turned out to be one of the best predictors for early DNA replication timing (Donaldson, 2005; Li et al., 2014). One good example is the differential replication timing of the tandemly arranged ribosomal genes (rDNA) of mice. 60% of these genes are replicated early in S-phase when they are transcriptionally active, demethylated and hyperacetylated. The remaining 40%, instead, are late replicating, transcriptionally inactive and marked by increased cytosine methylation and histone hypoacetylation (Li et al., 2005). This differences in replication timing are further manifested in nuclear location, with early replicating rDNA genes located in the nuclear and nucleolar interior, whilst the late replicating fraction is associated with the nuclear periphery .

Rtt109, which is responsible for acetylation of lysine 56 of histone H3 (H3K56ac) in yeast, shows direct interaction with several replication factors (Han et al., 2007; Suter et al., 2004). Also, origin firing in yeast is delayed upon deacetylation of histone tails around replication origins by the histone deacetylase Rpd3. On the other hand, increasing the acetylation level of histones associated with known late replicating origins has been shown to advance their initiation timing (Vogelauer et al., 2002). And drug induced hyperacetylation of large chromosomal regions or even entire chromosomes has been shown to lead to advanced replication timing (Casas-Delucchi et al., 2011; Casas-Delucchi et al., 2012) underpinning

the importance of this epigenetic modification. Notably, advances in replication timing of constitutive heterochromatin were achieved independently of changes in DNA methylation levels (upon knock out of the DNA methyltransferase *DNMT1*) or the methylation status of lysine 9 of histone 3 (after loss of the histone methyltransferase *SUV39h1/2*), both repressive marks for heterochromatin. Rather, in all three cases the authors observed changes in chromatin compaction of the pericentromeric heterochromatin clusters, which again suggests chromatin accessibility as a barrier to early origin activation.

In the same context and with regard to the aforementioned relation of chromatin (loop) organization and DNA replication it is interesting to note that acetylation of cohesin itself has been shown to affect DNA replication directly (Terret et al., 2009). Given that cohesin binds at regular intervals to chromatin, preferentially at specific insulator and boundary elements, it readily represents an obstacle for molecular machineries, including the replisome. Indeed, it was shown that binding and sliding of cohesin along yeasts chromosomes interferes with RNA polymerase II-mediated transcription (Gullerova and Proudfoot, 2008; Lengronne et al., 2004). The observation that acetylation had a direct influence on replication fork speed led to the suggestion that this modification turned into a conformational change of cohesin that resulted in better replisome passing along DNA (Terret et al., 2009). Interestingly, the acetylation of cohesin is mediated by the clamp loader RFC, which also loads PCNA to replication origins (Bermudez et al., 2003). Strikingly, both addition of the acetyl group through RFC and even sister chromatid cohesion seem to depend on the presence of PCNA (Song et al., 2012).

Finally, fork slowing also lead to smaller than usual chromatin loops in the work of *Terret et al.,* a finding that is in line with a previous study (Courbet et al., 2008). Here, the authors showed that differences in replication fork speed directly modulated loop sizes as well as the density of activated origins and changed the replication timing of the affected regions in the subsequent S-phase. Although no direct connection to the *timing decision point* during G1-phase could be shown, these findings provide a direct link for chromatin loop organization and DNA replication timing.

Chromosomal tandem repeat elements

During the last fifty years and even more with the advent of genome sequencing projects, it has been realized that repetitive DNA accounts for an unexpectedly large portion of genomes. Although repetitive sequences are usually underrepresented in sequencing studies, it has been estimated that more than 50 % of the human genome and up to 40% of mouse genomes consist of repetitive DNA which contains both mobile and tandem repeat elements, the latter referred to as *satellite DNA*. Perhaps due to the fact that such repeat elements have been difficult to study, they were long referred to as '*junk'* or '*parasitic'* DNA (Shapiro and von Sternberg, 2005). In stark contrast, only 1.2 % and 1.4 % of sequences in the human and mouse genome, respectively, account for protein coding sequences (Lander et al., 2001; Mouse Genome Sequencing et al., 2002).

Nowadays, it is beyond question that repetitive DNA plays an important role in the structural organization and functional regulation of the genome, and with that, maintenance of the integrity of every cell. At the simplest level, short nucleotide motifs serve as binding sites for various transcription factors which allows long-range interactions in *cis* to regulate gene expression.

Telomeric repeats are found at the end of linear chromosomes of most metazoans and are composed mostly of arrays of the double-stranded and non-coding $(TTAGGG)_n$ repeat element (Moyzis et al., 1988). Only the 3' end of telomeres is longer and single stranded, forming a so called T-loop structure (Griffith et al., 1999; Wright et al., 1997). Binding of the protein complex *shelterin* prevents repair of telomere structures by the DNA double-strand break repair machinery, due to the fact that they actually represent open and unprocessed DNA double-strand breaks. In addition to its protective function, *shelterin* also regulates telomere length in combination with the protein telomerase, which is only expressed in embryonic stem cells and many cancer cells (de Lange, 2005; Stewart and Weinberg, 2006).

In somatic cells, telomeres represent a difficult obstacle during DNA replication, as DNA polymerases are unable to fully duplicate the linear ends of chromosome, thus leading to telomere shortening with each cell division (Levy et al., 1992). This phenomenon is generally referred to as the *end replication problem* and leads to replicative and cellular senescence when telomeres have reached a critical length (Olovnikov, 1973). Mixed results have been obtained regarding the replication timing of chromosome ends, with significant differences between different species. In yeast for example, telomere replication takes

place late in S-phase and shifts to earlier replication with increased telomere shortening (Bianchi and Shore, 2007; Gilson and Geli, 2007). In mammalian cells, however, no relation between telomere length and replication timing has been observed and telomeres usually replicate throughout S-phase (Wright et al., 1997; Zou et al., 2004). Instead, nuclear positioning seems to affect the replication timing of individual chromosome ends, with earlier replicating telomeres positioned towards the nuclear interior (Arnoult et al., 2010).

Another particular class of genetic tandem repeats, the satellite DNA elements, make up a large fraction of mammalian heterochromatin. The two most prominent examples in mice are the centromeric minor satellite (MiSat) and the pericentromeric major satellite (MaSat) repeats. Major satellites are composed of 50 - 200,000 copies of a 234 bp-long AT-rich repeats that are arranged in arrays of up to 6 Mbp and which form large clusters in interphase nuclei, called *chromocenters* (Almouzni and Probst, 2011; Jorgensen et al., 2007). Minor satellite repeats (Kipling et al., 1991), on the other hand, are organized in ~600 kb domains of a 120 bp repeat unit.

Both satellite repeats share common epigenetic signatures such as increased DNA and histone methylation, histone hypoacetylation. However, they also exhibit differences in their chromatin composition, such as incorporation of a special histone variant (CENP-A at MiSat repeats) or the binding of three isoforms of heterochromatin protein 1 (HP1 at MaSat), which might contribute to their different functions (Guenatri et al., 2004). The observed clustering of chromocenters in mouse cells is stable over most of the cell cycle phases and become only separated with whole chromosome separation during mitosis (Quivy et al., 2004). Nonetheless, cell type and development specific differences in heterochromatin organization exist (Mayer et al., 2005).

In higher eukaryotes, centromeres form the massive kinetochore complexes, which form attachment sites for the mitotic spindle apparatus during mitosis and with that ensure proper sister-chromatid separation. Thus, they help in maintaining correct propagation of all genetic material, making them crucial chromosomal components (Guenatri et al., 2004).

Heterochromatin, including pericentromeric heterochromatin, is generally considered to be late replicating. Interestingly, studies in yeast and *Drosophila* indicated, that the replication centromeric DNA occurs early in S-phase and separated from the duplication of pericentromeric satellite DNA. Work performed in mice showed mixed results, where centromeres replicated either late in or throughout S-phase (Guenatri et al., 2004; Weidtkamp-Peters et al., 2006). The latter study, partially supports the findings from from

yeast and flies, that replication of centromeric and the surrounding pericentromeric DNA might be uncoupled, at least in some cases. These and other published and often contradicting results demonstrate that late replication timing is not strictly a marker of and also not a requirement for proper function of heterochromatin (Ahmad and Henikoff, 2001; Kim and Huberman, 2001; McCarroll and Fangman, 1988; Vig, 1995).

Aim of this study

The regulation of DNA replication timing in higher mammalian organisms is just as complex as the organisms themselves. This regulation likely relies on a precisely balanced interplay between many different factors rather than a sole determinant which dictates the initiation and further progression of DNA replication. Possible candidates are numerous and include, among others, epigenetic modifications like histone acetylation and DNA methylation as well as the architectural organization of chromatin in the cell nucleus.

Studies of the specific DNA replication programs at different developmental stages in *Xenopus* and *Drosophila* revealed unexpected and dramatic differences compared to somatic cells at later time points at the end of embryonic development. In mammals on the other hand, the replication program of somatic mammalian cells has been extensively characterized. However, much less is known about DNA replication dynamics in embryonic stem cells.

In this study, I aim to dissect the spatio-temporal DNA replication program of murine embryonic stem cells. Recent genome-wide studies have reported massive changes in replication timing during lineage commitment. I'm interested to know how these changes manifest *in situ*.

With the use of *fluorescence in situ hybridization* (FISH) I want to complement with information about the replication timing of important regulatory genetic elements like telomeres as well as heterochromatic tandem repeat elements, which are missing in those genome wide studies.

I further aim to compare the structural organization of chromatin and the functional unit of DNA replication - the replicon - in embryonic stem cells with that of somatic cells, published in a recent study.

Last but not least, I address the question if and how the loss of proteins responsible for the establishment and maintenance of global DNA methylation may affect DNA replication.

RESULTS

Characterization of the spatio-temporal DNA replication patterns in mESC reveals differences to somatic cells

With the aim to extend our knowledge on the organisation of DNA replication (Chagin et al., 2016) to pluripotent cells, I first performed a detailed characterization of the temporal order and frequency of DNA replication patterns in mESC as well as related measurements, such as population doubling time and S-phase duration.

Replication patterns are a direct visual representation of the spatial organization and temporal order of DNA replication. Thus, to characterize the DNA replication program in individual embryonic stem cells I used the nucleotide analogue EdU (*5-ethynyl-2'-deoxyuridine*) in combination with immunofluorescent detection of PCNA. The modified nucleotide is incorporated by DNA polymerases into newly synthesized DNA in cells that are undergoing S-phase. Thereby, sites of DNA synthesis are marked *in situ*. PCNA, on the other hand, as a central component of the replication machinery, localizes to active sites of DNA replication. These labeling approaches, allowed me to initially visualize and discriminate a total of five different replication patterns, including but not limited to those previously described in adult somatic cells (Fig. 4).

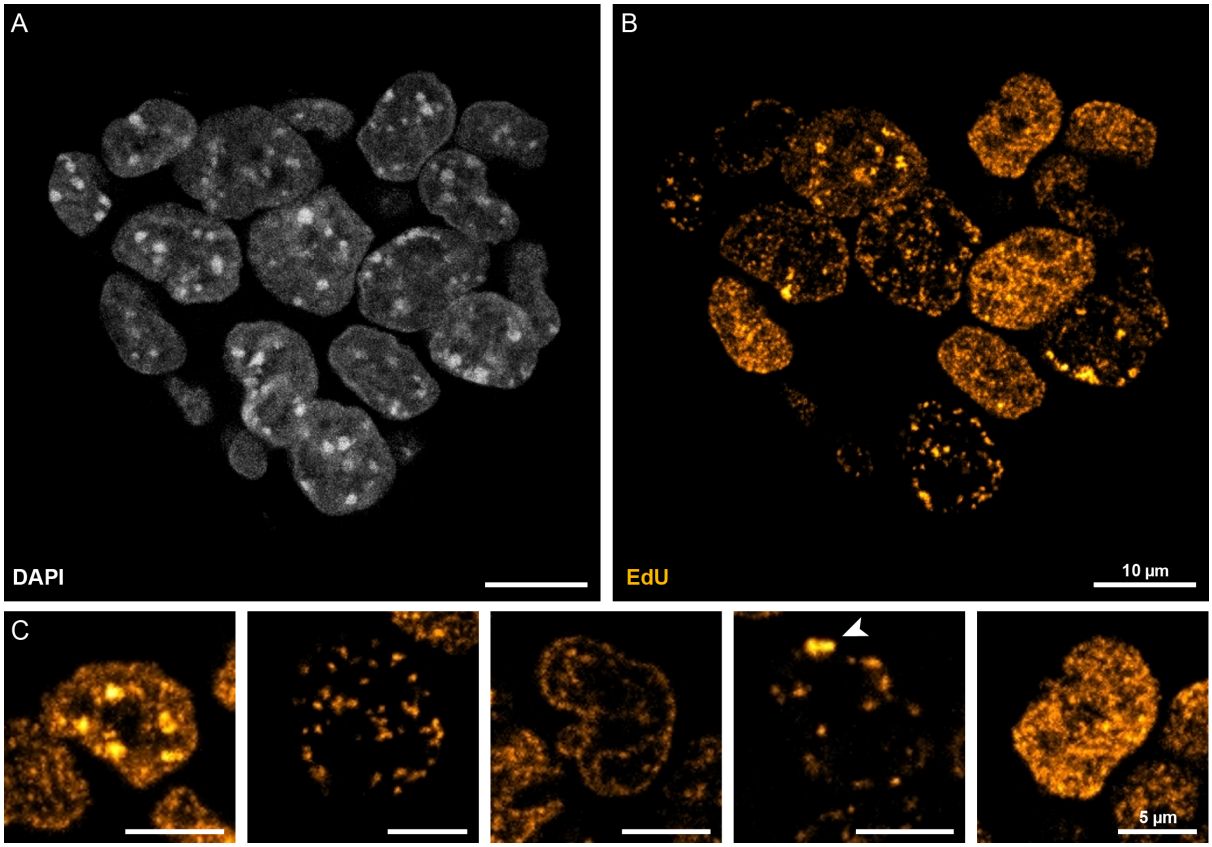


Figure 4 - Nucleotide labeling of newly synthesized DNA reveals five distinct replication patterns in murine embryonic stem cells. (**A**) Murine embryonic stem cells grow as 3D-spherical colonies in culture. Shown are mES cell nuclei counterstained with the DNA dye DAPI. Clusters of pericentromeric heterochromatin (chromocenters) become visible as bright spots. Nuclei are tightly packed within a colony, reflecting strong cell-cell interactions. Scale bar = 10 µm (**B**) Replication signals visualized in the cells shown in (**A**) upon incubation of the cells in medium containing EdU which labels nascent DNA. Scale bar = 10 µm. (**C**) Overview of the replication patterns observed in murine ES cells after nucleotide labeling. Based on the spatial distribution of the replication signal, a total of five clearly distinct patterns were frequently observed in mESCs. The white arrowhead marks a prominent structure which discriminates this pattern from the one shown in the second image in (**C**). Scale bar = 5 µm.

To unequivocally address the temporal order of the different patterns I then labeled mESCs as described above and performed different chase periods at 30 min increments between EdU labeling and fixation. EdU marks the sites of DNA replicated at the time of nucleotide application, while PCNA represents a later point during S-phase at the time of fixation of the cells. The use of these two replication markers and the increasing chase time, enabled us to discriminate two consecutive replication patterns at each increment to determine their temporal order as they become visible and more separated over time.

It further allowed me to identify cells that had either been in G1-phase at the time of nucleotide application or that entered G2-phase after DNA duplication had been completed. In the former case, the cell of interest would be negative for EdU (no incorporation during G1-phase) and stain positive for endogenous PCNA as soon as it entered S-phase. Like

this, the initial replication pattern that marks the beginning of S-phase, could be unambiguously identified. In the latter case, the cells are EdU-positive representing the last pattern at the end of S-phase. The PCNA signal, on the other hand, is expected to distribute within the G2-nucleus as it dissociated from the DNA and remains dispersed within the nuclear volume (Leonhardt et al., 2000).

I found that, at the beginning of S-phase, replication foci distributed homogeneously in the nuclear interior, reflecting duplication of the euchromatic portion of the genome as seen in other cell types (Fig.5A - I; (Leonhardt et al., 2000)).

As the signals became more separated with increasing chase period, large clusters of constitutive pericentromeric heterochromatin (termed *chromocenters,* (Mayer et al., 2005)) were replicated. In mouse cells, these structures can be visualized by counterstaining of the DNA with the fluorescent dye *4',6-diamidino-2-phenylindole* (DAPI) as densely stained clusters (Fig.5A - II). This observation was unexpected, given that pericentromeric heterochromatin is usually replicated latest in differentiated cells.

Replication of facultative heterochromatin is marked by the distribution of replication signals along the nuclear and nucleolar periphery. In mESC, I found a similar pattern at around the middle of S-phase (Fig.5A - III) which indicates that the replication timing of this chromatin type is conserved in pluripotent and somatic cells.

Pattern IV is characterized by replication foci that still locate mostly to the nuclear periphery. However, they decreased in number while increasing in size, suggesting some clustering of the underlying chromatin (Fig.5A - IV), similar to the clustered signals seen at chromocenters, although to a lesser extent.

Finally, the end of S-phase is marked by a strong accumulation of replication signal, limited to one particular region at the periphery of the stem cell nucleus (Fig.5A - V). In the next section, I will discuss in detail which structure is marked by this particular signal distribution.

The general characterization of the spatial distribution of replication signals in S-phase nuclei of embryonic stem cells led to the identification of a total of five replication patterns. In addition, I was able to also define their temporal order during S-phase. These findings are summarized schematically in Fig. 5B.

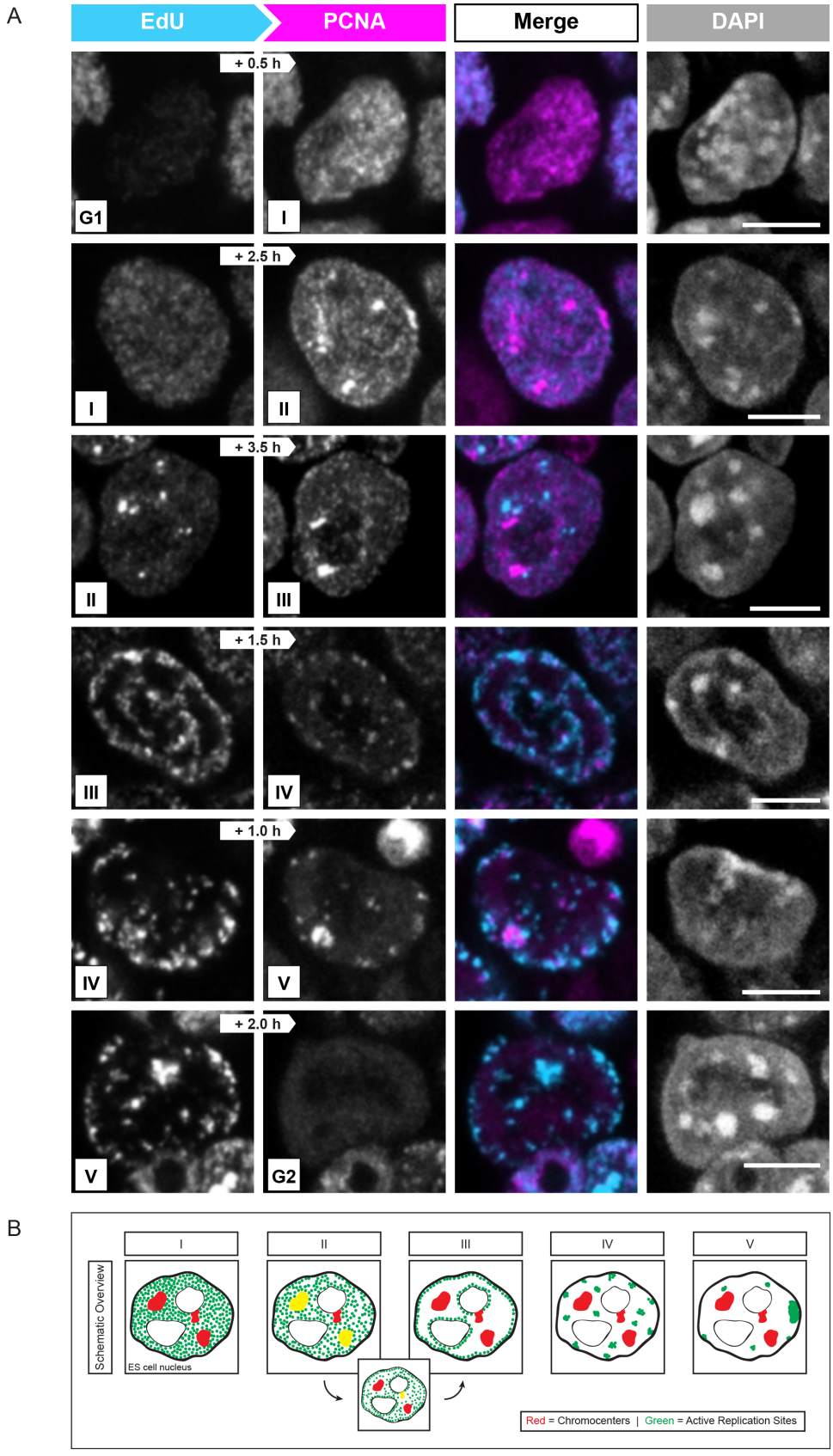


Figure 5 - Pulse-chase experiments reveal the temporal order of DNA replication in ES cells. (**A**) Fluorescent detection of the nascent DNA (cyan) together with endogenous PCNA (magenta) after different chase times (see white arrows in images) reveals the transition from one replication pattern to the next (from left to right in each row), representing the spatio-temporal order of DNA replication. In addition, a merge of the two

replication signals and DAPI-counterstained nuclei are shown. Top row: Cells negative for EdU are in G1-phase ([G1], no nucleotide incorporation). The first replication pattern that becomes visible after a short chase period (0.5 h) marks the beginning of S-phase [**I**]. Second row: pericentromeric heterochromatin gets replicated early in S-phase [**II**] and is followed by the duplication of facultative heterochromatin in the middle of S-phase, which is located at the nuclear and nucleolar periphery [**III**, third row]. The overlap of patterns **II** and **III** in the third row can be explained by the dynamic transition from one stage to the next. With increasing progression, the replication signals remain mostly at the nuclear periphery and decrease in number while increasing in size [**IV**]. The prominent structure shown in *Fig. 4* (white arrowhead) marks the end of S-phase [fifth row, **V**], after which the cells enter G2-phase [last row, **G2**]. Scale bar = 5 µm. **(B)** Schematic summary of the five replication patterns seen in mES cells as identified in (**A**). Replication signals are shown in green and pericentromeric heterochromatin in red. Replication of chromocenters (yellow due to overlap of "red" chromocenters and "green" replication signals) marks the second stage of S-phase but partially overlaps with the duplication of facultative heterochromatin at the third stage (inset between **II** and **III**).

To test if this progression through S-phase is a general feature of pluripotent stem cells I performed similar pulse-chase experiments in another commonly used ES cell line. E14 ES cells showed the same set and temporal order of the replication patterns (Fig. S1).

Finally, to test our findings *in vivo*, I performed live cell imaging experiments of J1 mESCs transfected with a plasmid encoding mCherry-tagged PCNA for up to 24 hours. To specifically assess the replication timing of chromocenters I co-transfected an eGFP-tagged major satellite-specific zinc finger binding protein (MaSat-GFP), that binds to pericentromeric regions. Fig. 6 shows the nucleus of a representative ES cell undergoing DNA replication. Notably, replication of the larger clusters of heterochromatin takes place in the first half of S-phase, as marked by the accumulation of mCh-PCNA at eGFP-MaSat spots (see *Stage II* at 1 hour 30 minutes). In addition, all other features described in fixed cells are reproduced *in vivo*.

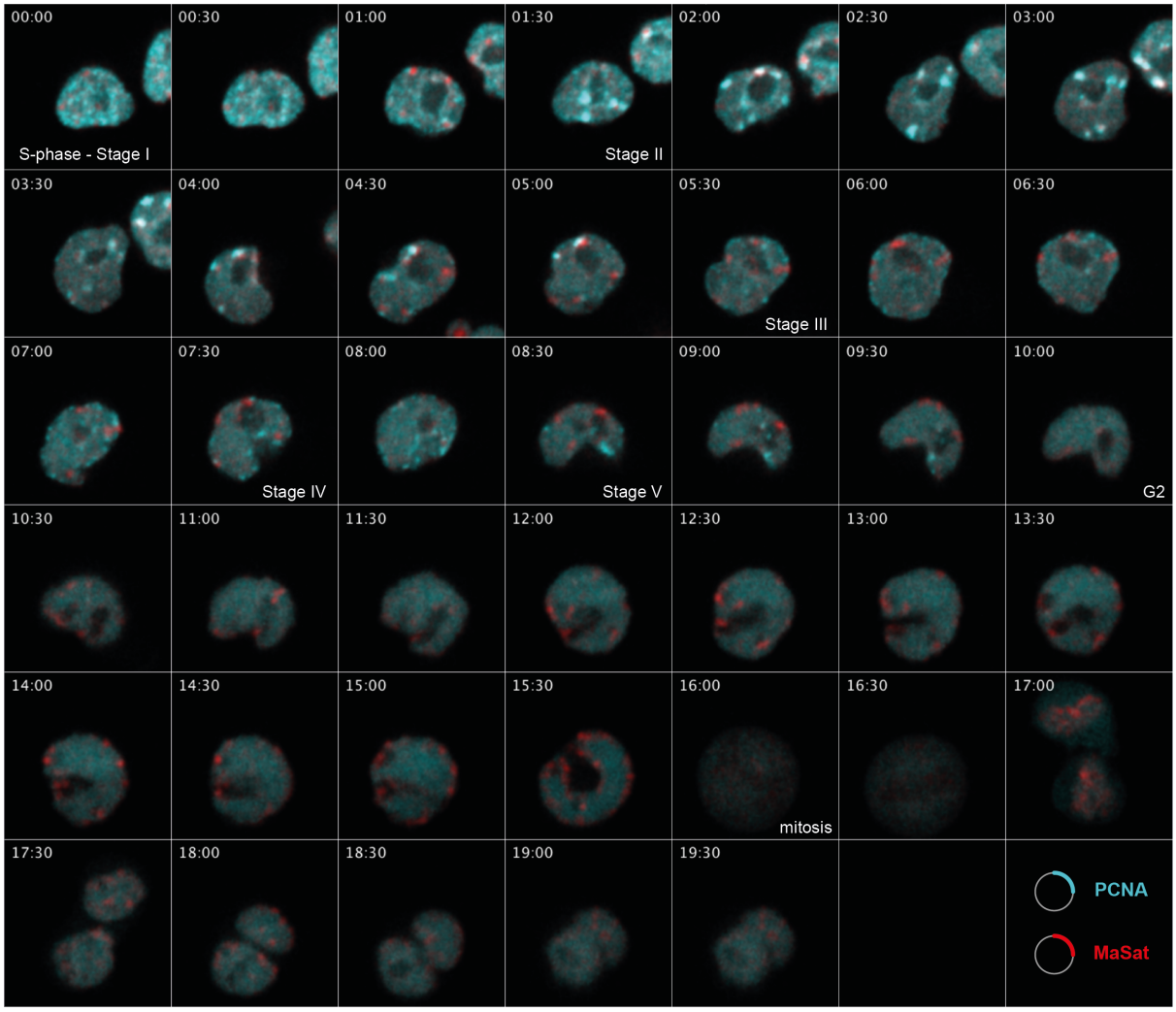


Figure 6 – Live cell imaging experiments confirm the DNA replication dynamics of embryonic stem cells. ***in vivo*** Representative ES cell transfected with plasmids encoding a mCherry-PCNA fusion construct (cyan) as well as a GFP-tagged major satellite binding protein (MaSat-GFP, red) that binds to and marks pericentromeric heterochromatin (chromocenters). The latter was used to specifically measure the replication timing of chromocenters. All replication patterns and their temporal order described in *Fig. 5* were reproduced *in vivo* (Stage I - V). After around 10 hours the cell enters G2-phase and division into two daughter cells during mitosis occurs after 16 hours.

A full understanding of the DNA replication dynamics of a specific cell type includes also knowledge about the population doubling time from which the duration of the different cell cycle phases (G1-, S- and G2-phase followed by mitosis) can be derived.

Embryonic stem cells are a special cell type in many respects. Besides their potential to indefinitely grow in culture and to differentiate into cell types of all three germ layer, they have been shown to possess a remarkably different cell cycle distribution when compared to somatic cell types (Li et al., 2012). Upon differentiation, cell cycle dynamics are

massively reorganized, which suggests that the characteristic cell cycle behavior of stem cells might be linked to and even required for the maintenance of the pluripotent state (Burdon et al., 2002; Soufi and Dalton, 2016; White and Dalton, 2005).

Their retained replicative capacity is one of the most remarkable features of ESCs and most cells are usually found in S-phase. This can be attributed to a shortened G1-phase with an otherwise S-phase duration that is comparable to that of somatic cells (White and Dalton, 2005). A consequence of this, is the rapid rate of cell division of only 8 - 10 hours (Stead et al., 2002). Interestingly, such a cell cycle distribution reminds of the early embryonic cleavage cycles of *Xenopus* and *Drosophila,* which lack detectable gap phases (Ferrell et al., 1991; Newport and Kirschner, 1982; Newport and Kirschner, 1984).

To confirm that the embryonic stem cell line retained this proliferative capacity and can be used in subsequent experiments, I measured the population doubling time of asynchronously growing J1 stem cell cultures over a time course of ten days (Fig. S2). Our measurements revealed a consistent doubling time of ~12 hours which is in agreement with doubling times published for other embryonic stem cell lines (Pauklin et al., 2011).

Next, FACS profiles of PI stained unsynchronized stem cell populations (Fig. 7A, quadruplicate experiments) were directly compared with differentiated cells (C2C12 mouse myoblast and HeLa cervix carcinoma cells (Fig. 7B), further confirming that most ES cells were in S-phase. However, without additional staining for cell cycle specific markers, FACS profiling of cells stained with quantitative DNA dyes allows only a relatively crude separation of subpopulations of cells into the different cell cycle phases (usually restricted to G1-, S- and G2-phase). A more detailed division into substages of S-phase beyond early or late S-phase, is thus not possible and consequently limits the temporal resolution that can be obtained from continuous DNA content profiles.

To overcome these limitations, I identified the fraction of replicating cells within fixed asynchronous cell populations by the quantification of cells that stained positively for replication signals (EdU). In the two samples analyzed (total N = 900 cells from 39 ES cell colonies) more than 70% of cells were found in S-phase, thereby confirming the results obtained by FACS profiling (Fig. 7C). Based on this frequency I calculated the approximate S-phase duration by multiplication with the cell cycle duration of 12 hours. This calculation reveals a S-phase duration of around 9 hours in J1 embryonic stem cells (Fig. 7D).

In a similar manner, I calculated the lengths of the individual sub-phases by multiplying the S-phase duration with the frequencies of cells at the different substages quantified in samples of the initial pulse-chase experiments (Fig. 7E). The stages were attributed based on the EdU signal (first pulse) and according to the replication patterns defined in the beginning.

This analysis confirmed a short G1-phase of around 2.5 hours. The approximately nine hours of S-phase are subdivided into the five different substages of which the duplication of euchromatin (pattern I) and pericentromeric heterochromatin (pattern II) takes most of the time (altogether ~6 hours). The last three substages are comparably short, with stages IV and V accounting for less than one hour before the cells finish S-phase and enter into G2-phase.

The pluripotent state of embryonic stem cells is maintained by the expression of a complex network of developmentally regulated transcription factors, including Oct4, Nanog, Sox2, and Klf4 (Masui et al., 2007; Mitsui et al., 2003; Niwa, 2007; Takahashi and Yamanaka, 2006). Besides classical *in vivo* test, like teratoma formation or the generation of chimeric mice, immunodetection of either of these markers are a valid method to assess pluripotency. I confirmed the expression of Oct4 in the ES cells used in this work via antibody staining, thus showing that they retained their pluripotent stem cell characteristics (Fig. 7F).

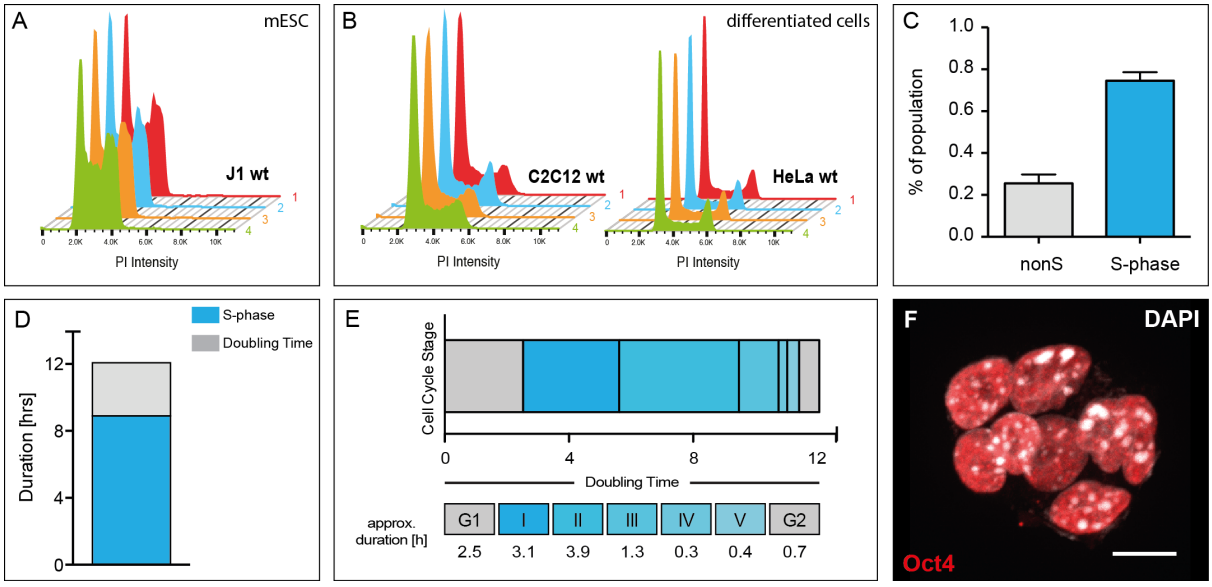


Figure 7 - Characterisation of the cell cycle distribution of murine ESCs. (**A**) Cell cycle profiles of propidium-iodide (PI) stained mES cells (quadruplicate experiments) compared to equally treated C2C12 mouse myoblast and human HeLa cells (**B**) to emphasize differences in the cell cycle distribution between pluripotent

and somatic cells. (**C**) Quantification of EdU positive cells within asynchronous ESC populations (N = 900 cells) confirmed that almost 80% of cells are in S-phase (mean ± SD). The population doubling time of J1 ES cells was determined by growth curve analysis over a course of 10 days (Fig. S2) and was around 12 hours. Together with the percentage of replicating cells from (**C**), this allowed approximation of the S-phase duration of about 9 hours for the male ES cells (**D**). In the same manner, the duration of the individual substages of S-phase was calculated from the fraction of cells at each stage (**E**). Immunofluorescent detection of the pluripotency marker Oct4 revealed clear accumulation of Oct4 in the cell nuclei and confirms that the ES cells are pluripotent (**F**). Scale bar = 10 µm.

Based on the presented results, with focus on the cell cycle characteristics of the embryonic stem cell line used in this study, I conclude that it is representative for pluripotent cells and suitable for use in further experiments. Interestingly and in contrast to a study published earlier (Panning and Gilbert, 2005), the data obtained from pulse-chase and *in vivo* live cell experiments suggest a DNA replication timing program that differs substantially between pluripotent ES and differentiated mammalian cells, as judged by the unusual sequence of replication patterns.

The Y-chromosome replicates in a synchronous manner similar to the inactive X-chromosome and marks the end of S-phase

In female cells, the inactive X-chromosome (Xi) replicates in a highly synchronous manner and, in contrast to the active homologue, within a short time interval during mid S-phase, as described previously by (Casas-Delucchi et al., 2011) (see also Fig. 3).

The end of S-phase in ES cells was marked by a prominent structure, at which replication signals accumulated in a manner, similar to that observed for the Xi (compare cell in midS in Fig. 3 and ES cell at stage **V** in Fig. 5). Given that the cell line used in our study (J1) was originally derived from a male mouse strain (129/terSv), I reasoned that this structure could be the male sex chromosome (Y-chromosome).

To test this hypothesis, I used *fluorescence in-situ hybridization* (FISH), which allows the detection of specific DNA sequences by hybridization with a complementary probe that contains fluorescently labeled or hapten-tagged nucleotide analogues (Cremer et al., 2008). I first performed FISH on both 3-dimensional structurally preserved J1 mESCs with a probe specific for the Y-chromosome. To assess the colocalization of the FISH probe with the

replication signal I co-stained the cells with an antibody specific for endogenous PCNA and observed clear colocalization between both signals (Fig. 8A).

Given that PCNA is a crucial component of the eukaryotic replisome and marks sites that are actively undergoing DNA replication, I measured the overlap of the Y-chromosome hybridization signal with that of co-stained endogenous PCNA in mESC and a male somatic cell line (mouse embryonic fibroblasts, MEF W8), a method we termed Repli-FISH (*Weber et al.*, under revision). This allowed us to assess if the late replication timing of the male sex chromosome is conserved during development. First, individual ES cells were categorized into the five substages of S-phase, based on the different replication patterns. In mouse embryonic fibroblast cells, four different patterns are observed that mark the progression through S-phase. These are depicted in Fig. 8C (middle row). I then computationally segmented the FISH probe signals in individual cell nuclei and finally measured the mean PCNA fluorescence intensities within each segmented signal and compared this measurement for the different S-phase substage (Fig. 8 B & C).

In both the embryonic stem and somatic cells, the strongest overlap was observed with the pattern representing the end of S-phase (**V** in mESC and **IV** in MEFs, respectively), clearly demonstrating that the Y-chromosome is the last structure that gets replicated during the S-phase and that this replication timing is conserved during development.

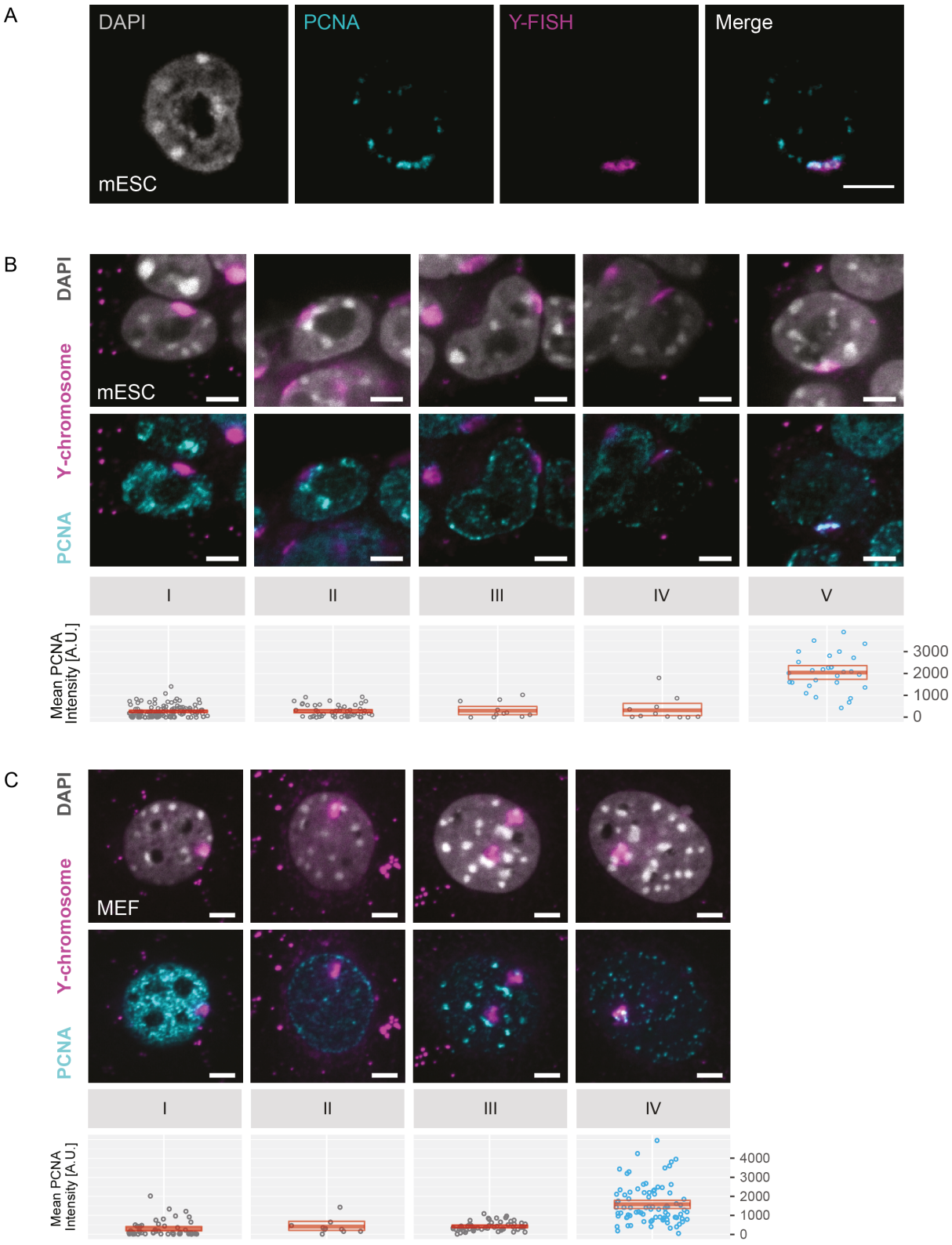


Figure 8 – Replication of the Y-chromosome marks the end of S-phase in male cells. The synchronously replicated structure at the end of S-phase (*Fig. 5*, **V**) corresponds to the Y-chromosome as revealed by the clear overlap of the replication signal with Y-specific FISH probe (**A**). Scale bar = 5 µm. Colocalization analysis of the Y-chromosome and the replication signal over the whole S-phase in male ES (**B**) and mouse embryonic fibroblast cells (MEF W8, **C**) supported this finding. Cells hybridized with the Y-chromosome specific probe were analyzed at all S-phase. Top rows in (**B** and **C**) show the localization of the Y-chromosome in the cell nucleus.

Identification of replication patterns was based on the PCNA signal distribution (middle row). Plots show the mean PCNA intensity within the FISH signal (each dot represents one Y-chromosome). Red boxes represent mean value (bold line in the middle) and 95% confidence intervals. Timing of replication during S-phase of the Y-chromosome is reflected by increasing PCNA signal at the Y-chromosome and was highest at the last substage of S-phase in both cell lines (**B** and **C**). Scale bar = 5 µm

In the study of *Casas-Delucchi et al.,* replication of the (peri)centromeric heterochromatin of the inactive X-chromosome was unsynchronized from the rest of the chromosome and instead took place together with that of the other chromosomes late in S-phase (Casas-Delucchi et al., 2011). In the Repli-FISH samples analyzed, chromocenters that could be visually related to the Y-FISH signal in both cell lines, also appeared to replicate earlier, i.e. together with pericentromeric heterochromatin from other chromosomes (see stage II in mESC and stage III in MEFs in Fig. 8 B & C, respectively).

It would be interesting to investigate how and why the replication timing of pericentromeric regions on both chromosomes differs from that of the remaining chromatin. Given that at least the inactive X-chromosome is marked by a specific combination of repressive epigenetic modifications (H3K27m3, accumulation of *Xist* RNA, histone hypoacetylation and increased DNA methylation) it could be possible that the observed difference in replication timing can be attributed to the different set of repressive marks that are present on pericentromeric heterochromatin (H3K9m3, increased DNA methylation, binding of the heterochromatin protein HP1).

In summary, I conclude that duplication of the Y-chromosome occurs synchronously at the end of S-phase in male cells throughout development and that replication of pericentromeric satellite DNA is uncoupled from the remaining chromatin of the male sex chromosome.

Analysis of the replication timing of chromosomal tandem repeat elements

Noncoding DNA makes up an unanticipated large fraction of many eukaryotic genomes (de Koning et al., 2011) and is no longer considered ‘junk DNA’. Such sequences, including repetitive elements, carry out diverse and important regulatory functions in mammalian cells, such as the organization of chromatin in the nucleus with the help of *matrix associated regions* (MARs). Tandem repeat sequences may also act as binding platforms for factors involved in chromatin regulation and, to stay in the context of DNA replication,

short non-coding sequence motifs serve as determination sites for replication initiation, at least in prokaryotic and lower eukaryotic organisms (Shapiro and von Sternberg, 2005). Another prominent example for the role of repetitive sequences in transcriptional regulation is a phenomenon called *position effect variegation*. When placed in close linear proximity to pericentromeric heterochromatin of chromosomes in yeast or *Drosophila,* genes that are usually active become repressed (Weiler and Wakimoto, 1995).

Besides their importance for cellular metabolism, genetic repeats also represent difficult obstacles for the molecular machineries acting on them. In the context of DNA replication, for example, the multitude of repetitive sequences, telomeres and also centromeres, are difficult to access by the factors required for DNA synthesis and, as a result, are hard to replicate (Aze et al., 2016; Miller et al., 2006; Sfeir et al., 2009; Usdin et al., 2015). Detailed knowledge about these elements and their functions can thus help to understand their mechanistic role e.g. in transcriptional regulation of genes or proper chromosome segregation during mitosis.

In genome-wide experiments, however, repeat elements are difficult to map. This fact hampers the measurement and detailed study of their epigenetic landscapes, their structural organization within the nucleus and their replication-timing from data obtained in similar studies. To expand our understanding of the replication timing of three major chromosomal tandem repeat elements, i.e. centromeric DNA (*minor satellite repeats* in mice, MiSat), peri-centromeric DNA (*major satellite repeats* in mice, MaSat) and telomeric repeats in embryonic stem cells, I again performed Repli-FISH experiments on 3D preserved stem cell colonies with PCR generated probes for each of the repeat elements (see *Materials & Methods* and Tab. 7).

The specificity of each probe for its intended target sequence has been confirmed in previous studies (Anton et al., 2014; Frauer et al., 2011; Lehnertz et al., 2003), and is further supported by the localization of the respective signal on mitotic chromosomes (shown for MiSat and telomeres in Fig. 9).

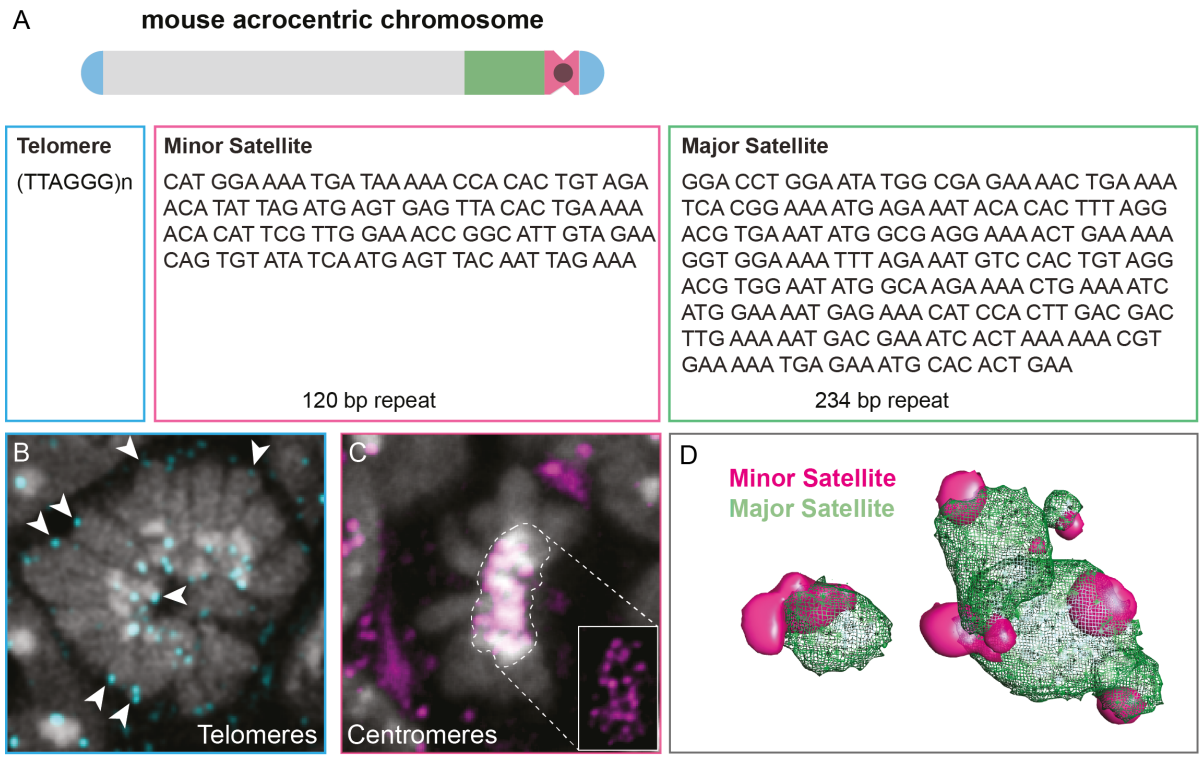


Figure 9 – Description of the tandem repeat elements analyzed in embryonic stem cells. (**A**) sequence motif of different tandem repeat elements (blue = telomeres, magenta = centromeric heterochromatin, green = pericentromeric heterochromatin) and their organization along a schematic mouse acrocentric chromosome. The specificity of minor satellite and telomere probes could be confirmed by their distribution on mitotic chromosomes (**B** and **C**). Telomeres are seen as punctate signals at the end of each chromosome (white arrowheads), while centromere signals localize only to regions of higher DAPI intensity representing centromeric and pericentromeric heterochromatin. (**D**) In line with their linear organization along chromosomes, minor and major satellite signals organize as spatially related structures in interphase cell nuclei as shown by a 3-dimensional reconstruction of both signals (minor satellite signal based on FISH probe, major satellite signal based on DAPI counterstaining).

Telomeres cap both ends of the individual chromosomes in order to protect them from degradation and fusion after each round of genome duplication (Arnoult et al., 2010). Minor satellite repeats, on the other hand, form the functional `centers` of acrocentric mouse chromosomes by providing an architectural framework for the association with cohesin during mitosis to form the kinetochores (Pidoux and Allshire, 2005). Major satellite repeats are associated with centromeres and surround the centromeric repeats on each chromosome (see Fig. 9A). Pericentromeric heterochromatin is known to repress transcription of transposable elements and might be involved in the general organization and maintenance of heterochromatic regions (Almouzni and Probst, 2011)

When visualized with repeat-specific FISH probes in interphase nuclei, minor satellite signals are seen as small focal structures located in close proximity to chromocenters, the latter can be identified by bright DAPI counterstaining (Fig. 10B and (Guenatri et al., 2004)) as confirmed by FISH (Fig. 10A). This close spatial association is also shown in a 3D-reconstruction of minor and major satellite signals (Fig. 9D). Although diploid mouse cells contain a set of 40 acrocentric chromosomes, clustering of these chromosomal regions, probably mediated through interaction of specific heterochromatin binding factors like HP1 (*heterochromatin protein 1*; (Guenatri et al., 2004)), results in a lower than expected number of signals upon visualization. Visual inspection and manual counting of major and minors satellite FISH signals in mESCs support the idea of clustering of these regions as indicated by the low average numbers of specific signals for each repeat element (~17 and 27 for major and minor satellite probes, respectively (see Fig. S3). In the case of telomeres, the FISH signals distribute as even smaller individual foci throughout the cell nucleus (Fig. 10C) and clustering into larger aggregates was less apparent (average number of signals ~58).

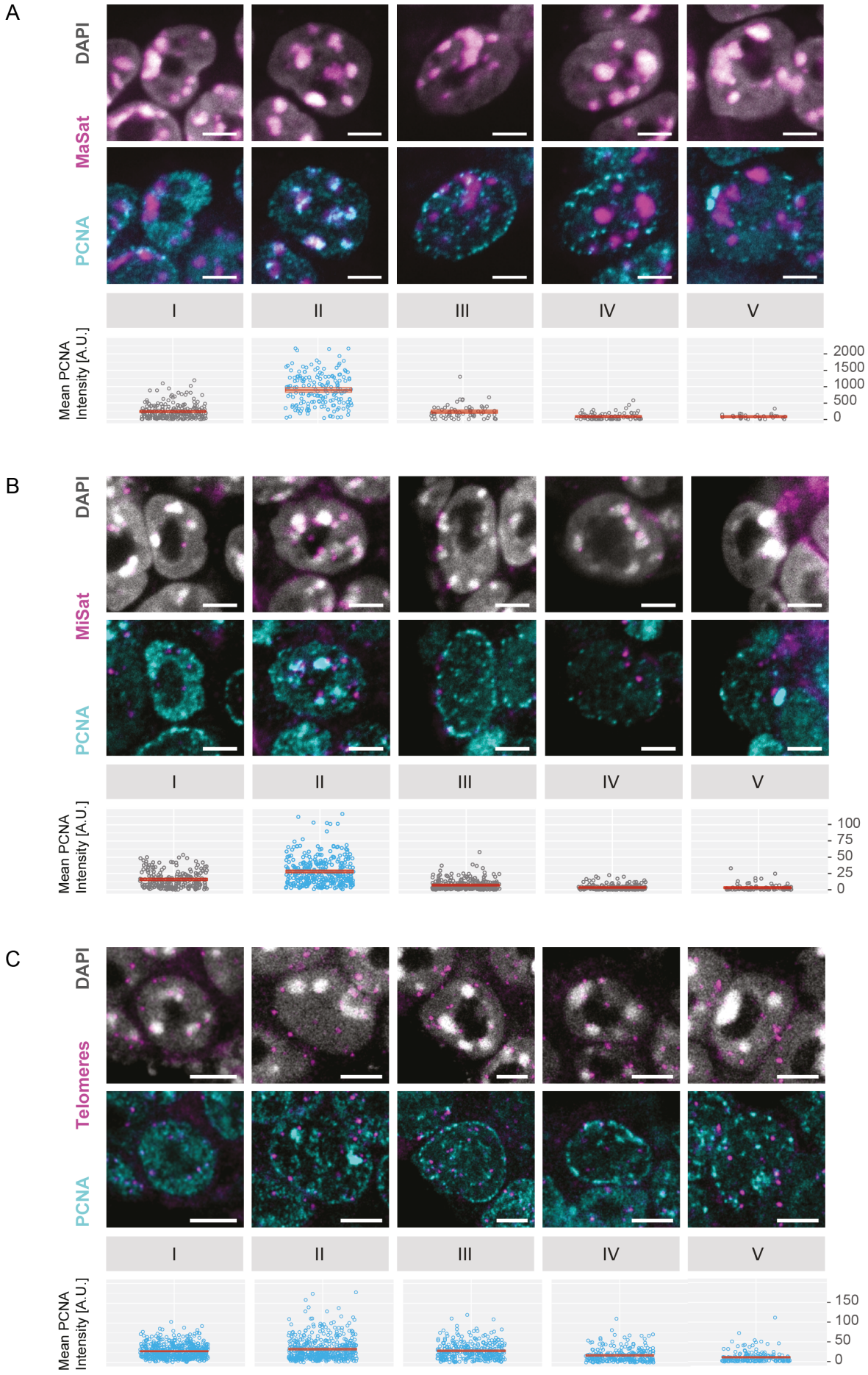

A
DAPI
MaSat
PCNA
I
II
III
IV
V
Mean PCNA Intensity [A.U.]
2000
1500
1000
500
0
B
DAPI
MiSat
PCNA
I
II
III
IV
V
Mean PCNA Intensity [A.U.]
100
75
50
25
0
C
DAPI
Telomeres
PCNA
I
II
III
IV
V
Mean PCNA Intensity [A.U.]
150
100
50
0

Figure 10 – Analysis of the replication timing of chromosomal tandem repeat elements in mESCs. The specific nuclear localization of tandem repeats can be seen in the overlays of interphase nuclei and FISH signals (top rows in **A - C**). Replication timing analysis of tandem repeat elements was done by Repli-FISH as in Fig. 8 (see Fig. S4 for further details). Accumulation of PCNA at both pericentromeric (major satellite; **A**) and centromeric repeat DNA (minor satellite; **B**) is strongest at S-phase substage II and reflects duplication of these repeats in the first half of S-phase (top and middle panel). Telomeres (bottom panel), on the other hand, are replicated throughout S-phase (**C**), as reported before for human telomeres in somatic cells (Wright et al., 1997; Zou et al., 2004). Red boxes represent mean value (bold line in the middle) and 95% confidence intervals. Scale bar = 5 µm

The results presented in Fig. 10 A-C represent the overlap of PCNA with individually segmented FISH signals within the interphase nuclei of stem cells. The strongest overlap of both major and minor satellite signals with PCNA occurs during stage II of S-phase (Fig. 10 A & B). This indicates that not only pericentromeric, but also centromeric chromatin is preferentially duplicated during this early stage of S-phase. Previous studies addressing the replication timing of the two satellite repeat classes in two mouse fibroblast cell lines (NIH 3T3 (Guenatri et al., 2004) and MEFs (Weidtkamp-Peters et al., 2006)), presented different results for the duplication of minor satellite repeats. In the first study, they were found to replicate asynchronously from major satellite repeats in late S-phase, while the work of *Weidtkamp-Peters et al.* suggested that replication occurs at all stages of S-phase. It is possible, however, that the replication timing of centromeric regions is coupled to or influenced by the early replication of the closely associated pericentromeric heterochromatin in ES cells.

The overlap of PCNA and telomeric repeat signals, on the other hand, was comparable throughout the different stages and dropped only towards the end of S-phase (Fig. 10C). This suggests that replication of telomeres occurs asynchronously throughout S-phase in mESCs. While in yeast cells the duplication of telomeres occurs late during DNA replication (Friedman et al., 1995), the replication of most human telomeric regions also occurs over the whole course of S-phase (Wright et al., 1999). I therefore conclude that the replication timing of telomeres could be more conserved in vertebrate organisms.

Given that these observations are in good agreement with published results, we additionally conclude that the Repli-FISH method is a suitable approach allowing sensitive measurement of the replication timing of specific genetic loci *in situ*, i.e. in interphase nuclei of mammalian cells.

The special organization of pericentromeric heterochromatin in embryonic stem cells might explain its advanced replication timing

Chromocenters are marked by the histone modification H3 trimethylated at lysine 9 (H3K9m3, (Peters et al., 2001)), are hypoacetylated (Jeppesen et al., 1992), exhibit increased levels of DNA methylation (Guenatri et al., 2004) and are bound by specific heterochromatin proteins like HP1, as mentioned before (Eissenberg and Elgin, 2000).

Previous studies showed that cell types of different origin or developmental status have a different organization of pericentromeric heterochromatin (Mayer et al., 2005; Solovei and Joffe, 2010; Solovei et al., 2009). Such differences in organization could be mediated, for example, by specific protein-protein interactions between heterochromatin binding factors. The MBD protein family member MeCP2, for example, was shown to be responsible for the clustering of pericentromeric heterochromatin during terminal differentiation of mouse myoblasts and plays an important role in the neurological disorder *Rett syndrome* (Agarwal et al., 2011; Agarwal et al., 2007; Brero et al., 2005). Furthermore, the activity of different chromatin assembly and remodelling . As such, knock-down of the chromatin remodelling complex CAF-1, which is also essential for the rapid cell division of *Xenopus* embryos, resulted in the disruption of the organization of pericentromeric heterochromatin specifically in murine ES cells but not mouse embryonic fibroblasts (Houlard et al., 2006).

I reasoned that a different degree of chromatin compaction might account for the differences in replication timing of pericentromeric heterochromatin in murine ES cells that were observed in the initial experiments. To assess differences in compaction, I compared the organization of chromocenters in the cell nuclei of mouse myoblasts and mouse embryonic stem cells, counterstained with DAPI and imaged by 3D-structured illumination microscopy. Fig. 11A shows two representative nuclei from a mouse myoblast (upper row) and a mouse ES cell (lower row).

I counted and compared the average number of chromocenters per nucleus (Fig. 11B) which reflects differences in clustering of the pericentromeric regions from multiple chromosomes. In addition, I compared the surface areas (Fig. 11D) and the shape factor of all chromocenters within the cell nuclei (Fig. 11C) as a measure for chromatin compaction. The shape factors reflects morphological deviations of a 3-dimensional structure from that of a 3D-sphere of the same volume as the object-of-interest (i.e. a chromocenter). A perfect sphere would result in a shape factor of 1, while smaller values indicate an increasing deviation from this globular structure (Fig. 11E).

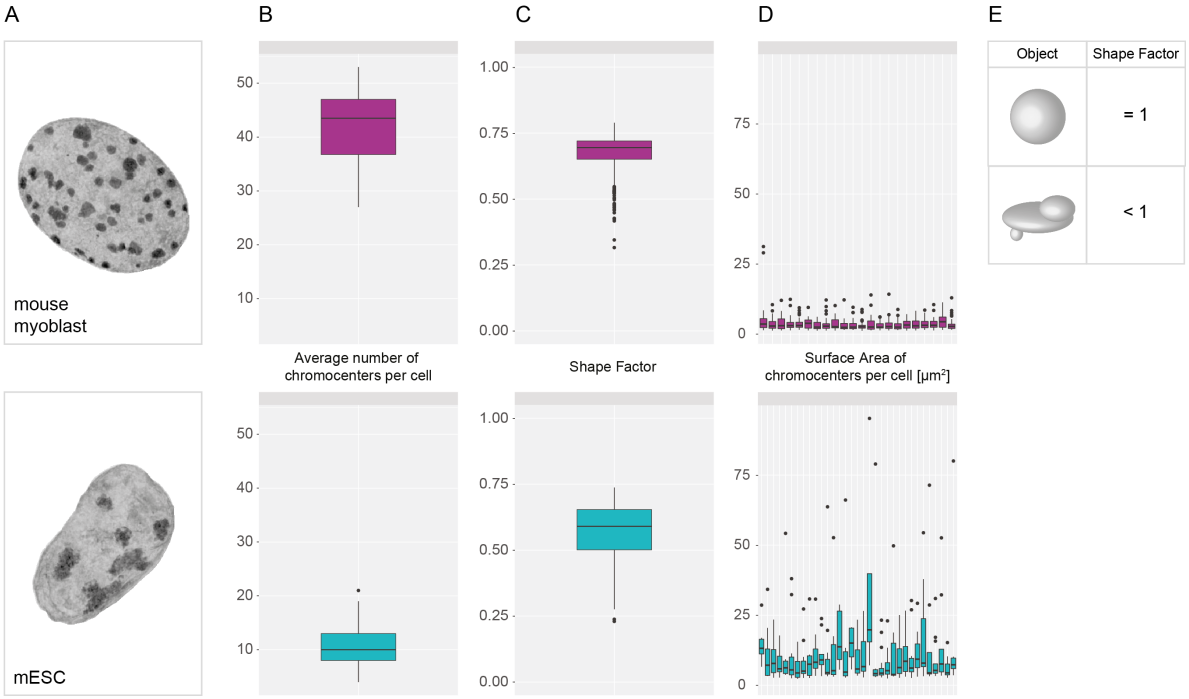


Figure 11 - Pericentromeric heterochromatin is less compact in embryonic stem cells which could explain its advanced replication timing. (**A**) Representative cell nuclei from somatic mouse myoblasts (upper row for all measurements) and embryonic stem cells (lower row). From these images, the different organization of pericentromeric heterochromatin (chromocenters) becomes apparent. Boxplots show median values, inter-quartile range, upper and lower Whiskers as well as potential outlier values (black dots; see also Fig. S6). Differences in clustering is reflected by the higher average number of chromocenters in the somatic cells (**B**) as well as a higher Shape Factor (**C**). This factor reflects the `roundness` of an object in comparison to a perfect 3-dimensional sphere of the same volume (Shape Factor = 1, (**E**)) and was used as a measure for chromocenter compaction. The diffuse organization of pericentromeric heterochromatin in ES cells results in larger surface areas with much higher variability (**D**, lower row). In all myoblasts, on the other hand, the surface areas of chromocenters were almost the same in all cells analyzed (**D**, upper row).

The differences in chromatin organization that occur during development and differentiation are in part reflected by the differences in clustering of chromocenter (Brero et al., 2005; Kobayakawa et al., 2007). As seen in Fig. 11 B, clustering is less prominent in C2C12 mouse myoblast cells, as they contain around four times more chromocenters than pluripotent cells. Even when accounting for the genome size differences between the two cell lines (quasi tetraploid (C2C12) versus diploid (mESC), the number of chromocenters in ES cells is still around 2-fold lower, which is consistent with the results of a previous study that directly compared undifferentiated and differentiated mESC (Kobayakawa et al., 2007).

The higher average shape factor obtained for all chromocenters of the myoblast cells (0.68 in C2C12 vs. 0.57 in mESC, respectively; **** with p > 2.2e-6 measured by unpaired two-

sample Wilcoxon test) further indicates that chromocenters of mouse myoblasts were more spherical and thus compacted. This is further supported by the almost equal distribution of the surface areas of chromocenters from multiple individual C2C12 cell nuclei (Fig. 11D). On the other hand, the pericentromeric clusters in ES cells showed a much higher variability with an overall larger average surface area (3.5 μm^2 for C2C12 vs 10.7 μm^2 in mESCs).

Although I did not directly investigate differences in, for example, histone acetylation levels or differences in the expression and/or binding of heterochromatin proteins between the two cell lines, the obvious differences in heterochromatin morphology and clustering within the cell nuclei and the increased surface areas of chromocenters in ES cells, indicate that the pericentromeric heterochromatin of murine embryonic stem cells is less compact. This is in agreement with results obtained in earlier studies (Kobayakawa et al., 2007; Meshorer et al., 2006) and provides a profound mechanistic basis for the observed differences in the DNA replication program.

Molecular properties of the stem cell replicon

Complete genome duplication once per cell cycle and in the restricted time frame of S-phase basically depends on only two determinants. These are i) the number and distribution of initiation sites (i.e. replication origins) along the genome and ii) the processivity of replication forks emanating bidirectionally from these sites.

Given that origins are not defined by DNA sequence in higher mammalian species and that the ORC complex also has no apparent sequence specificity, the choice of replication start sites occurs dynamically and needs to be re-established anew during each round of genome duplication (Machida et al., 2005). This situation appears to be even more drastic during early development.

Indeed, the spacing of adjacent origins within replicon clusters in *Xenopus* and *Drosophila* is highly dynamic, resulting in significantly different replicon sizes compared to later developmental stages. This is reflected e.g. by the much shorter *inter-origin distances* (IOD) measured in *Xenopus* nuclei before the *mid-blastula transition* (MBT; 5-15 kb, (Blow et al., 2001; Hyrien and Mechali, 1993)) compared to later stages after the MBT (150 - 300 kb; (Hyrien et al., 1995)). Replicon sizes appear to be not only flexible in the context of development, but can be also highly variable within the same cell population.

Moreover, *Conti et al.* showed that in response to the variations in inter-origin distances in isolated human keratinocytes, also their replication fork speed was altered (Conti et al., 2007) and the mouse myoblasts used in (Chagin et al., 2016) relied on an increase in replication fork speed to duplicate their tetraploid genome rather than the activation of additional replication origins. This suggests that the dynamic control of origin spacing, origin activation and with that replicon sizes as well as variations in the speed of replisomes are tightly linked to each other and that varying these parameters provides a mechanisms for cells to adapt when time is a limiting factor. With regard to the limited availability of activating factors and nucleotides for the synthesis of new DNA, it is unclear how such changes are regulated.

With regard to the differences in their spatio-temporal replication dynamics uncovered in the beginning of this work, I was interested to know if mES cells exhibit further replication-related adaptations. I asked whether such adaptations could occur at the same level as the developmental variations seen in frogs, namely the organization and molecular properties of stem cell replicons. *Chagin et al* performed many of the experiments that are suitable to address these questions in mouse myoblast and human cancer cells and their measurements were representative for other published somatic cells (Chagin et al., 2016). I therefore related the results from the subsequent experiments to those of these two cell lines.

To analyze the molecular characteristics of replicons in ES cells, I made use of the molecular combing assay, which allows the measurement of replication fork speed and inter-origin distances from large numbers of individual DNA molecules isolated from cell populations (Bialic et al., 2015; Bianco et al., 2012). Asynchronous populations of the cells of interest are first and consecutively labeled with two differently halogenated nucleotide analogues for a desired amount of time. Like EdU, these analogues are incorporated into newly synthesized DNA during replication.

Upon isolation, digestion of associated chromatin proteins and stretching of high molecular weight genomic DNA on silanized glass surfaces thousands of replication tracks can be visualized on individual DNA fibers (Bianco et al., 2012; Conti et al., 2001; Techer et al., 2013). While pulling the surface modified coverslips out of the DNA solution, a receding air-water meniscus guarantees a constant and sequence independent extension of DNA fibers which results in a conversion factor of 1 µm = 2 kb of DNA. This procedure frequently results in average fiber lengths of 250 - 500 kbp (Lebofsky and Bensimon, 2003).

J1 embryonic stem cells were first labeled for 15 min with 5-iodo-2'-deoxyuridine (IdU), followed by a second pulse of the same length with 5-chloro-2'-deoxyurdine (CldU). I then extracted high-molecular genomic DNA and performed the molecular combing assay (Fig. 12 A).

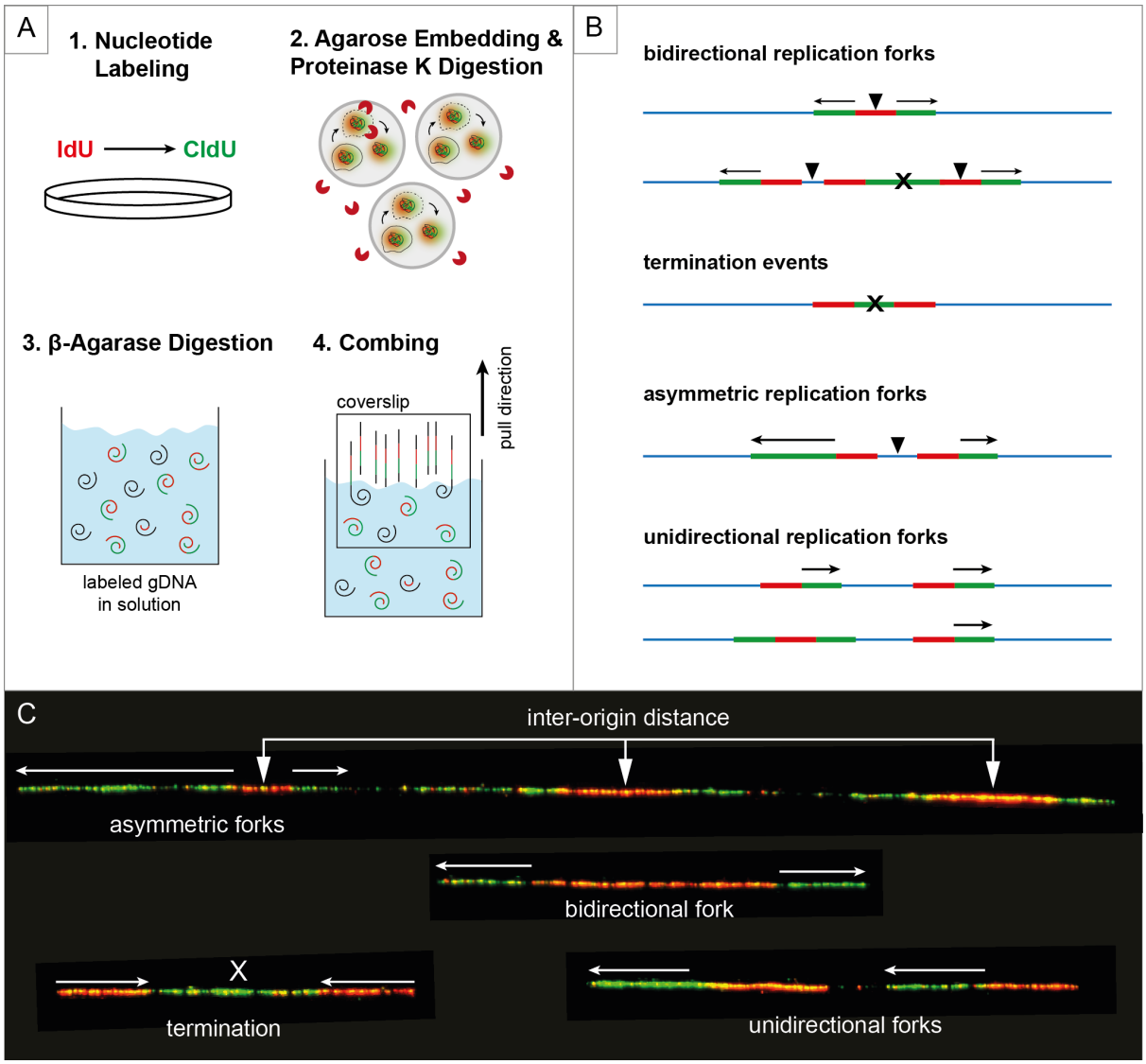


Figure 12 - Evaluation of the molecular parameters of embryonic stem cell replicons. (**A**) Schematic overview summarizing the Molecular Combing procedure. Asynchronous cell populations are consecutively labeled with two differently halogenated nucleotide analogues (red: IdU; green: CldU) for a desired amount of time (*1*). After that, cells are harvested by trypsinization and embedded in low-melting point agarose followed by digestion with proteinase K to remove cellular and chromatin proteins (*2*). Agarose plugs containing high molecular weight genomic DNA are digested with β-agarase to release DNA into solution (*3*). Finally, the DNA is combed on silanized glass coverslips. After an initial incubation to allow attachment of DNA fibers to the glass surface, the coverslips is pulled out of the DNA solution at a constant rate (*4*). This results in thousands of linear DNA molecules that can be analyzed microscopically. (**B**) Classification of typical signals obtained from Molecular Combing experiments. Replication fork speed is calculated from measurements of green replication

signal (arrows) flanking red tracks that mark initiation events during the first pulse labeling period (origins of replication marked by black arrowheads). Merging forks within clusters of origins and termination events are discarded (marked with X). Fork asymmetry can be assessed by calculating the ratio of left/right replication forks, emanating from a origin of replication. Replication signals spreading in only one direction are considered as unidirectional replication forks. (**C**) Representative examples of replication signals from DNA combing experiments. Labeling is according to the classification described in (**B**). The inter-origin distance (IOD) is the distance between neighboring initiation sites within a cluster, identified with the help of the first nucleotide label (IdU, red).

After measuring suitable signals (Fig. 12 B & C), I calculated a replication fork speed in mESCs of 1.69 ± 0.02 kb/min (mean ± SEM, Fig. 13A)**.** This result is within the range of fork rates measured for HeLa cells (Chagin et al., 2016) and of that published for other cell lines (Techer et al., 2013), indicating that no significant differences exist at the level of replication fork speed between pluripotent and somatic cells. The two forks that proceed in opposite directions from the same origin also appear to travel at almost the same rate as reflected by the measured fork asymmetry (Fig. 13 B), suggesting that the replication forks of embryonic stem cells do not experience a high level for stalling events due to exogenous or endogenous factors.

On the other hand, the average inter origin distance, i.e. the distance between two neighboring origins of replication, turned out to be shorter in the mES cells compared to the mouse myoblasts and human cells (~90 kb in mESC compared to 161.7 kb (C2C12) and 188.7 kb (HeLa), respectively, Fig. 13C).

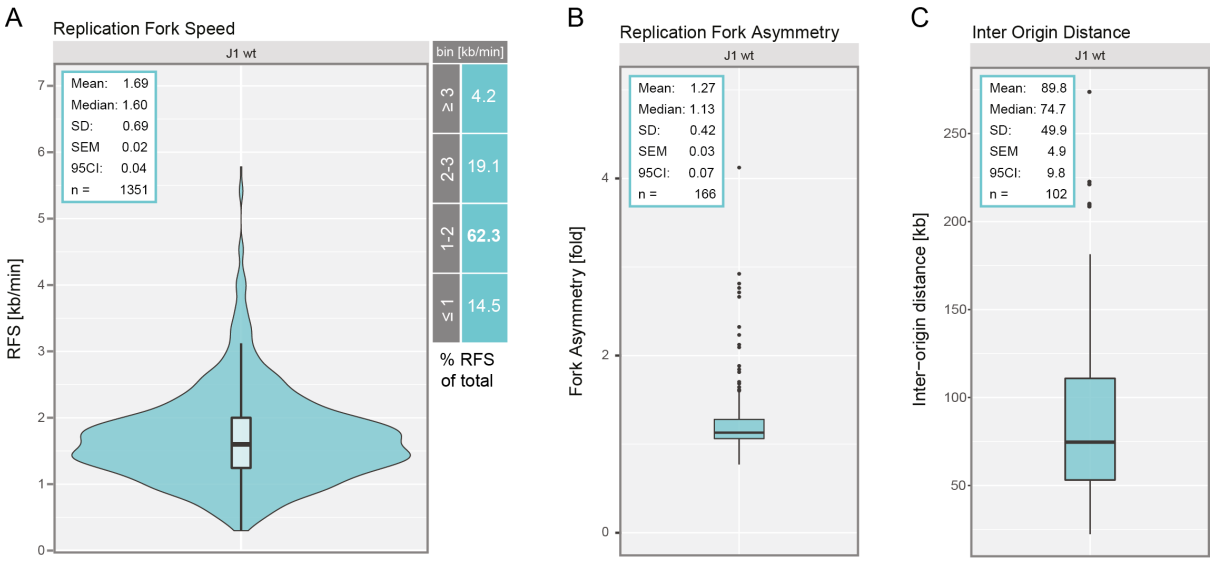


Figure 13 - Evaluation of the molecular parameters of embryonic stem cell replicons. The DNA fiber assay was used to assess molecular characteristics of an average replicon in mouse ES cells from large populations of individual DNA fibers. This included the measurement of replication fork speed (RFS; **A**), fork asymmetry (FA; **B**) and inter origin distances (IOD; **C;** (see Fig. S6 for description of boxplots and violinplots). Average replication fork speed of mESC was 1.69 ± 0.02 kb/min (mean ± SEM, N = 1351 measured replication tracks). Fork speeds measured in the range of 1-2 kb/min accounted for more than 60% of all measured signals, reflecting the robustness of the measured average RFS. Replication forks were not highly asymmetric (median ratio of left/right replication fork ~1.1, N = 166 measured bidirectional forks), which would otherwise have indicated differences in the speed of forks at one side of a bidirectional replication bubble. The average IOD in the pluripotent cells was ~90 kb (median 74.7, N = 102 measured distances), clearly shorter than that of differentiated mouse and human cells as reported earlier (not shown, 161.7 kb (C2C12) and 188.7 kb (HeLa), ;(Chagin et al., 2016)).

This results shows that the organization of replicons or replicon clusters is different between mouse pluripotent cells and somatic cells. Although the differences between mES and mouse myoblast or human cells are not as dramatic as those of the early and late stages of *Xenopus* embryos, this still suggests that the modulation of inter origin distances, and with that the resulting replicon sizes, represents a mechanism that is similar between the two species and which represents further developmental differences of the replication timing program of murine ES cells.

The measurement of these parameters provides important information for our understanding of the molecular dynamics of replication machineries in different cells and cell types, but not only that. They can also help to improve our understanding of the organization of DNA at the molecular scale and the potential influence of the topology of chromatin on the regulation of DNA replication.

3D-SIM replication nano-foci measurements in ES cells

In the last decade, advanced optical microscopy techniques have been developed to allow imaging beyond the resolution limit inherent to light microscopy. By using multicolor 3D-structured illumination microscopy (3D-SIM), it was recently shown that individual replication foci are resolved down to single replicons and, to some extent, even individual replication forks (Chagin et al., 2016).

This finding is incongruous with the idea of *replication factories* as synthesizing machinery that remain attached to a nuclear scaffold during DNA synthesis and that simultaneously duplicate multiple clusters of replicon (Hozak et al., 1994). The spatial proximity of replicons can instead be explained by the underlying chromatin conformation, rather than its physical interaction with fixed *replication factories.* Supported by earlier FRAP studies (Leonhardt et al., 2000; Sporbert et al., 2002), individual replicons appear to be processed by replisomes that are assembled *de novo* onto DNA during S-phase and spatially related origins, e.g. those of a replicon cluster, become activated in a domino-like manner over the course of S-phase (Casas-Delucchi and Cardoso, 2011; Leonhardt et al., 2000; Lob et al., 2016). Like this, the propagation of origin firing events could be maintained until the whole genome has been duplicated. Importantly, this mechanism would also allow functionality in *trans* and affect initiation of origins located on different chromosomes.

Given that the chromatin organization and chromosomal interactions are known to be different in embryonic stem cells and differentiated cell types (Dixon et al., 2015), I aimed to analyze whether the relation between replicons and replication foci imaged with 3D-SIM (hereafter referred to as nano replication foci or nanoRFi) remains the same during early mouse development. This could help to answer whether differences in the organization of the underlying chromatin fiber, e.g. towards a more open chromatin conformation, may have an impact on the organization of replication origins and their activation during S-phase, which could, for example, also explain the advanced replication timing of pericentromeric heterochromatin.

I quantified the number of super-resolved nanoRFi of mES cells the different substages of S-phase (patterns I-V) after growing them in the presence of BrdU for a short period of time (Fig. 14A). Segmentation and counting of the BrdU nanoRFi of individual cell nuclei was done according the protocol described in (Chagin et al., 2015).

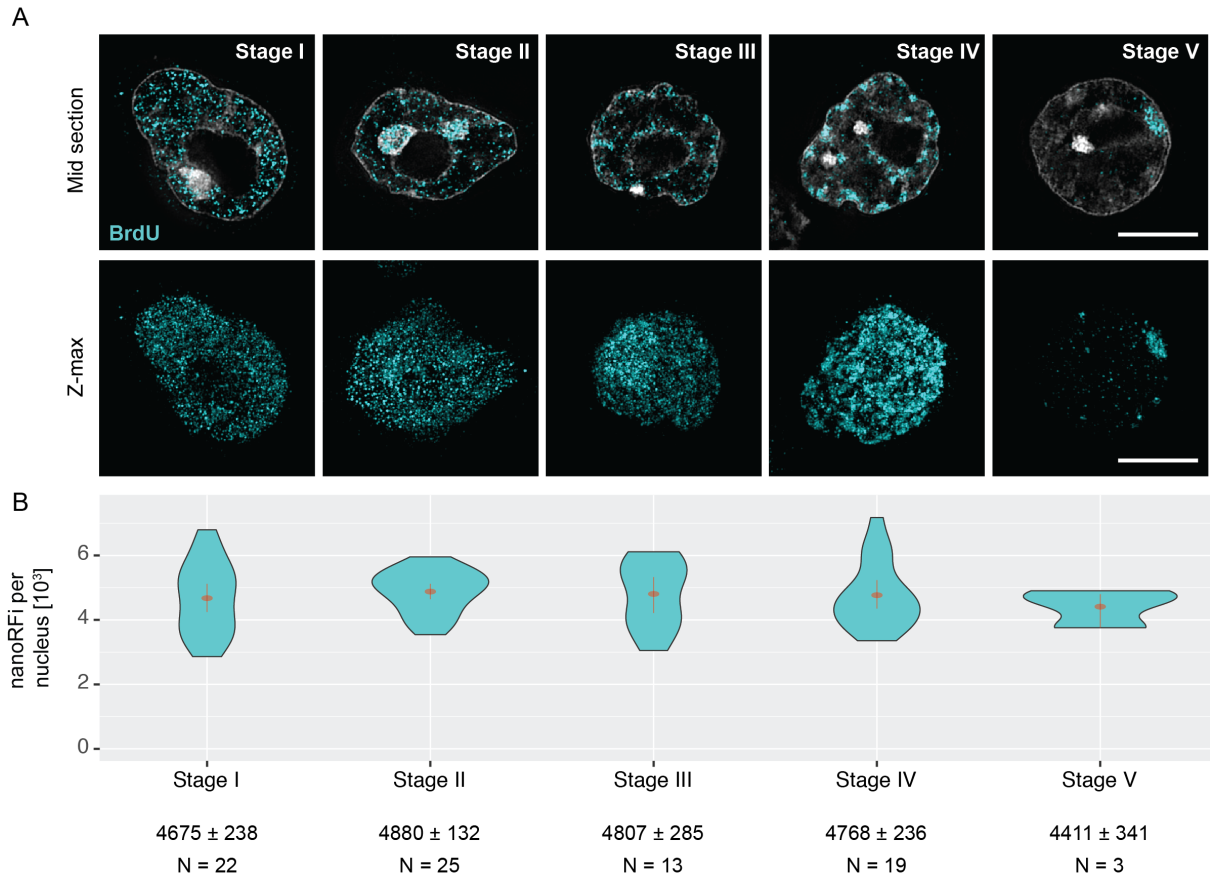


Figure 14 - Measurement of DNA replication signals by 3D structured illumination microscopy. (**A**) Mid focal sections (top row) of stem cells imaged by 3D-structured illumination microscopy. Nuclei (DAPI) at the different substages of S-phase, identified by the distribution of replication signals (BrdU) within the nucleus. Maximum projections of the replication foci of each cell are shown in the lower row. Scale bar = 5 µm. (**B**) Average numbers of super-resolved replication foci in embryonic stem cells revealed an average of 4700 replicating sites present at any given time during S-phase. Distribution of foci numbers for all cells are plotted as violin plots, separately for each substage. Numbers below represent mean ± SEM (see Fig. S6 for description violin plots and summary of the data in Supplementary Tab. 2). N = number of cells analyzed.

I obtained comparable numbers of nano replication foci for cells at all different S-phase substages (Fig 14 B and Tab. S2) with an average of 4708 ± 246 (mean ± SEM) foci present at any time during S-phase in embryonic stem cells. Interestingly, also the nanoRFi measurements for cells at stage II were not dramatically influenced by the more compact organization of pericentromeric heterochromatin (4880 ± 132). Clustering of the replication signals within chromocenters presented an obstacle for the segmentation procedure in the cell lines used in the previous study and were thus omitted from the analysis. The fact that this is not the case in mESCs could be explained by the differences in pericentromeric heterochromatin compaction in the ES cells (Fig. 11) and supports the idea that differences in the underlying organization of the DNA fiber somehow influences DNA replication (Lopes Novo and Rugg-Gunn, 2016; Meshorer and Misteli, 2006; Meshorer et al., 2006).

Our measurements revealed that approximately 4700 genomic sites are actively undergoing DNA replicating at any given time point during S-phase in mES cells, a number that is comparable to those measured in C2C12 (5314 ± 227) and HeLa cells (5583 ± 162). Both somatic cell lines exhibit a polyploid karyotype, thus it would be interesting to know if the existence of a comparable number of replication foci in ESCs is also due to a higher genome content.

Calculation of the genome size of mESCs

As discussed earlier, variations in the number and spacing of activated origins can serve as a mechanism to ensure genome duplication in the available time during S-phase in the light of karyotype changes in a cell, were the initiation of an increased number of start sites results in more active replication forks, acting together to accomplish the duplication of additional genetic material.

Most established cell lines used for *in vivo* studies have initially been transformed in order to immortalize them and to revive their replicative potential, e.g. upon transfection of the enzyme *human telomerase reverse transcriptase* (hTERT). Such procedures, however, often introduce undesired genetic aberrations which are known to drastically influence the ploidy of the transformed cell line. Embryonic stem cells, on the other hand, are derived from the inner cell mass (ICM) of a developing embryo and are, by nature, capable to sustain their proliferative state in culture. Moreover, they have been described to maintain a stable diploid karyotype (Martin, 1981).

To relate the high number of replication foci obtained in the previous measurements, I performed karyotype analysis of J1 embryonic stem cells from metaphase chromosome preparations. Using published genome data available from the *Genome Reference Consortium* (https://www.ncbi.nlm.nih.gov/grc/mouse/data) for haploid mouse genomes, I then calculated their approximate genome size in combination with the chromosome counts obtained from metaphase spreads.

Manual counting of >100 metaphase spreads confirmed a diploid karyotype of the mouse J1 cell line, consisting of 40 acrocentric chromosomes in mice (Fig. S5). To derive the total genome size in megabases we used the most current mouse genome assembly *GRCm38.p6* that provides sizes for each chromosome of a haploid mouse genome. From

this we calculated a diploid genome size of approximately 5.19 gigabasepairs (Gbp, see Table 1).

Table 1 - Genome size estimation based on karyotype analysis and published genome size data

Chromosome	Size (Mb)	Chromosome	Size (Mb)	Chromosome	Size (Mb)
1	195.47	**8**	129.4	**15**	104.04
2	182.11	**9**	124.6	**16**	98.21
3	160.04	**10**	130.7	**17**	94.99
4	156.51	**11**	122.08	**18**	90.7
5	151.84	**12**	120.13	**19**	61.43
6	149.74	**13**	120.42	**X**	171.03
7	145.44	**14**	124.9	**Y**	91.74
Σ = 2725.52 Mbp = ~2.73 Gbp (haploid) **= ~5.19 Gbp (diploid)**					

In contrast, the mouse myoblast and human cell lines exhibited a genome size of 11.4 and 9.7 Gbp, respectively. We note that our estimation excludes repetitive DNA elements which are not covered in whole genome sequencing techniques. However, these regions are estimated to not extend more than a few hundreds of megabase pairs (~300 - 400 Mbp, A. Rapp, personal communication). This result therefore shows that an equal number of replication foci is present in embryonic stem cells with only around half the genomic content than the somatic cell lines used by *Chagin et al.*

Embryonic stem cells appear to activate a high number of replicons

Having determined the genome size, S-phase duration and number of nano replication foci as well as molecular parameters of replicons in the ES cells, it is possible to calculate the total number of origins that become activated during S-phase and also to compare the organization of replicons, as the smallest unit of DNA replication, between embryonic and somatic cells. The relevant parameters are summarized in Tab. 2.

Table 2 - Comparison between important S-phase related parameters in pluripotent and somatic cells

Cell Type	S-phase duration [hours]	Genome Size (GS, in Gbp)	Replication Fork Speed (RFS, in Ntd/min)	Inter-origin distance (IOD, in kbp)	nanoRFi at any given time
mESC (J1)	9.3	5.19*	1690	89.8	4708
C2C12†	9.4	11.4	2460	161.7	5314
HeLa†	9.5	9.7	1650	188.7	5583

* estimation based on published genome data; † data taken from (Chagin et al., 2016)

The total number of replicons, reflecting the number of origins activated during S-phase, is given by the genome size divided by the average inter origin distance obtained from DNA fiber experiments (Tab. 3)

Table 3 - Calculation of the total number of replicons activated during S-phase in mESCs

Measurement	Formula	mESC (J1)	C2C12*	HeLa*
Total number of activated replicons	GS/IOD	57,714	70,501	51,404

*: data taken from (Chagin et al., 2016)

This calculation revealed that a comparable number of origins are activated during S-phase in all three cell types and irrespective of the differences in genome size. This result could indicate that the high number of replicons in the somatic mouse and human cells are a result of their increased genome content, requiring the activation of additional origins of replication over the course of S-phase. In the previous study, this may have remained unnoticed, given that both cell lines have a comparable genome size.

The measurement of inter origins distances can be influenced, however, by the sample quality and the length of fibers obtained in DNA combing experiments (Techer et al., 2013), representing a technical limitation of this method. As a result, origins that are far apart from each other cannot be properly identified and measured resulting in an inherent bias for IODs within a certain range, defined by the average fiber lengths of the sample. Thus, it remains possible that there exist differences in the organization of replicons between ES and somatic cells that are not properly detected solely by IOD measurements. However, to address this possibility additional experiments that could allow more sophisticated measurements of replicon size and organization are required.

As measurements of replication fork speed are less prone to the length of DNA fibers, I used the replication fork speed to determine the number of replicons active in parallel at any time during S-phase for comparison with somatic cells and in order to relate them with the number of measured 3D-SIM replication foci (Tab. 4).

The genome size divided by the average speed of a replication fork represents the time required to synthesize the entire genome if only a single replication fork would be active for the entire length of S-phase. Dividing this time by the actual measured S-phase duration, thus effectively represents the number of all active bidirectional replication forks at any given time during S-phase in the studied cell line. Since replicons, per definition, are the unit of DNA duplicated by a pair of bidirectional forks, half their average number represents the replicons that are active in parallel.

Table 4 - Evaluation of the relation between 3D-SIM nanoRFi and replicons in mESCs and somatic cells based on the parameters summarized in Tab. 2

Measurement	Formula	mESC (J1)	C2C12*	HeLa*
Time to replicate the genome with one fork	GS/RFS [**hours**]	51168.31	77235.77	97979.8
Replication forks active in parallel	GS/RFS/ S-phase duration	5491.32	8216.57	10313.66
Replicons active in parallel	active forks/2	2745.66	4108.29	5156.83
Replicons per nanoRF	calculated replicons active in parallel/counted nanoRFi	0.58	0.77	0.92

*: data taken from (Chagin et al., 2016)

Interestingly, this showed that in mESC, a lower number of replicons is active in parallel (2746 (mESC) versus 4108 (mouse myoblasts) and 5157 (HeLa) active replicons, respectively). This further suggests that the larger genome size of the somatic cells used in the work of *Chagin et al*, required the simultaneous activation of more origins of replication, which ultimately results in a comparable number of total activated replicons in all three cell types to achieve full genome duplication in a comparable length of S-phase.

Regarding the relation between replicons and the replication signals imaged by super-resolution microscopy it was shown that each of the nano replication foci in somatic corresponds to, on average, one replicon (Fig. 15 A - C and Tab 4 - *Replicons per nanoRF*; (Chagin et al., 2016)). The lower ratio obtained for mouse myoblasts could be explained by differences in the clustering of replication signals or a higher frequency of nanoRFi that contained single replication forks that could be resolved by 3D-SIM. The ratio of replicons per replication focus in embryonic stem cells was even a bit lower and in fact correlates better with the number of individual replication forks (Fig. 15 C).

In principle, this discrepancy could be explained in two ways, schematically represented in Fig. 15 D). i) the number of unidirectional replication forks is much higher in embryonic stem cells than in other cell types which would affect the calculation of replicons active in parallel (Tab. 4) or ii) the combination of a less compacted chromatin in ESCs (Kobayakawa et al., 2007; Meshorer et al., 2006) and the resolving power of the 3D-SIM system allows to measure even more individual replication forks rather than replicons.

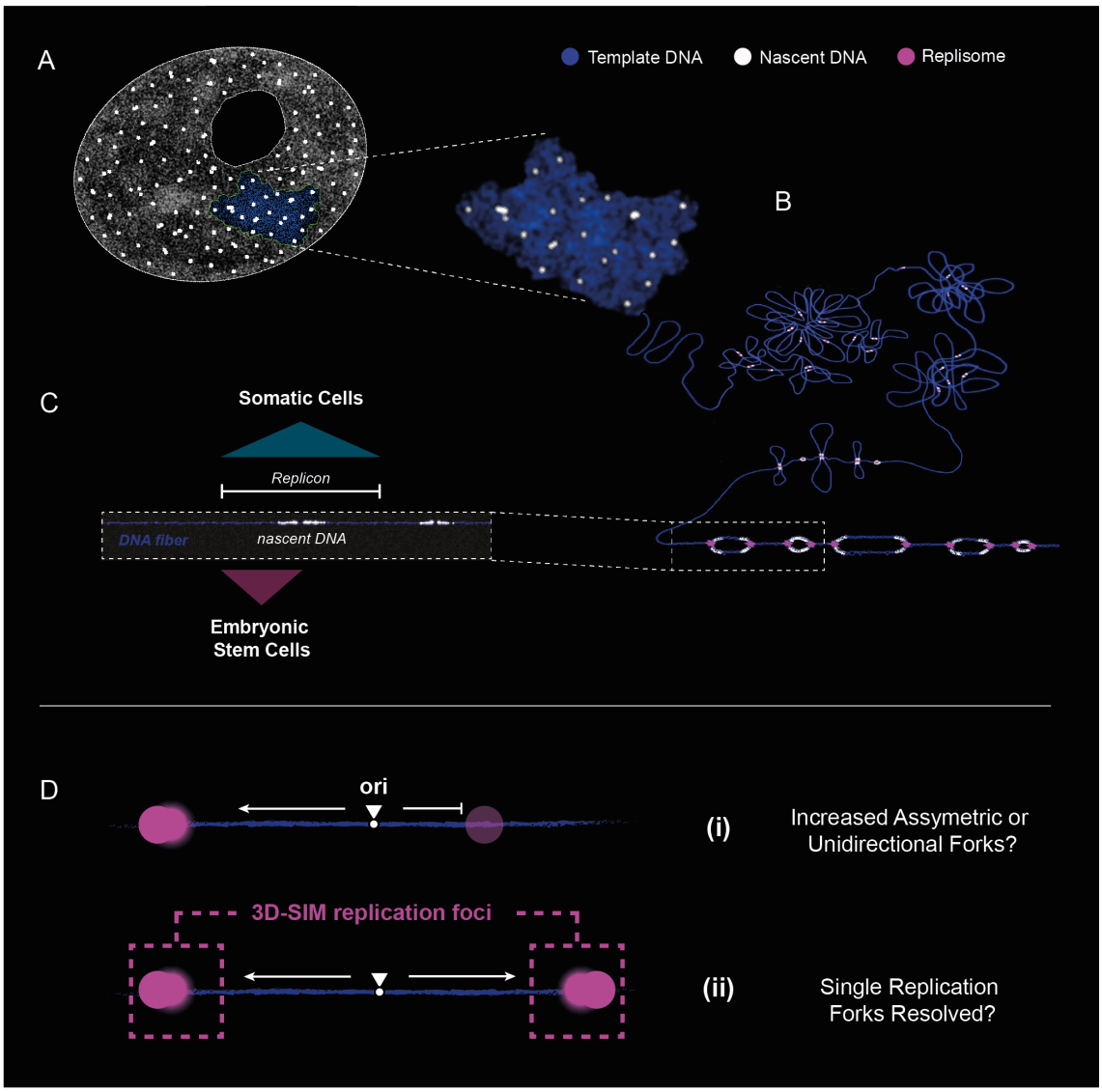


Figure 15 - Relationship between replicons and replication nano foci in embryonic stem cells. During DNA replication, nascent DNA becomes visible as discrete foci, represented as white spots in a schematic representation of an interphase cell nucleus and an enlarged chromatin region (**A**). Following the different hierarchical levels of chromatin in the cell nucleus, each focus was shown to correspond to individual replicons in somatic cells, e.g. as the ones seen in DNA fiber experiments (B & C, (Chagin et al., 2016)). Measurement and comparison of calculated replicons and nanoRFi in embryonic stem cells, however, suggest that one replication nano focus corresponds to only around half a replicon, or individual replication forks (C). (D) This discrepancy could be explained in two ways: (i) In addition to bidirectional replication forks, embryonic stem cells possess a higher number of unidirectional or asymmetric forks that would influence the calculation of active replicons (see Tab. 4) or (ii) the fraction of individually resolved replication forks (instead of replicons) is much higher than in somatic cells, probably due to a less compact chromatin in mESCs combined with the resolving power of the 3D-SIM system.

As it is well-established that DNA replication generally occurs bidirectionally, one-sided replication forks are expected to represent only a minor fraction of all forks. I therefore assessed the percentage of unidirectional forks from the Molecular Combing data and found it to be around 5% of all the replication forks considered suitable for our RFS measurements (data not shown). This is comparable to the frequency published for mouse embryonic fibroblast cells in a recent study (Stanojcic et al., 2016). However, it is important to note that this number was subject to a high variation. Also, due to the variability of signals seen in single-molecule data and the limitations imposed by typical fiber length obtained in such experiments it is possible that the frequency of unidirectional forks is underestimated.

One-sided replication forks could also be a result of a multitude of endogenous and exogenous factors that are known causes of fork stalling events. Such replicative stress can either alter DNA directly (like the generation of single- or double-strand breaks by intrinsic repair pathways or ionizing radiation), or else, DNA metabolic processes as a result of base modifications that present physical barriers to the molecular machineries (Zeman and Cimprich, 2014). As described earlier, we also did not measure a significant asymmetry of bidirectional forks in the single-molecule data (see Fig. 13 B) and under physiological conditions, i.e. in the absence of replicative stress, no biologically relevant function of unidirectional replication forks has been envisioned so far. Thus, from the available data that can be used to assess both unidirectionality or stalling of replication forks, we conclude that neither of the two options are the cause of the observed discrepancy between replicons and nanoRF in ES cells.

To test if, indeed, higher rates of single replication forks are resolved in mESCs, I pulse labeled C2C12, HeLa and J1 mESCs for 20 minutes with the nucleotide analogue BrdU and immediately fixed them after this incubation time. Given that a combination of differences in chromatin organization and the better optical resolution could account for the discrepancy between replicons and nanoRFi, the structural organization of the underlying DNA should possibly remain conserved.

PCNA represents a central component of the eukaryotic replisome (Boehm et al., 2016; Sporbert et al., 2005). I therefore argued that the measurement and comparison of the fluorescence intensities of immunostained endogenous PCNA within individual nano replication foci may be used as a readout for single replication forks per focus. Consequently, a higher fluorescence intensity of PCNA in somatic cells would reflect more, i.e. eventually two, forks within one focus. In turn, if replicons in embryonic stem cells

contained mostly single forks, I could expect to measure a lower average fluorescence intensities of PCNA within BrdU nano-foci.

Super-resolved RFi from BrdU as well as PCNA were segmented as described earlier and the total fluorescence intensity of PCNA strictly overlapping with BrdU foci was measured and plotted for all three cell lines (Fig. 16).

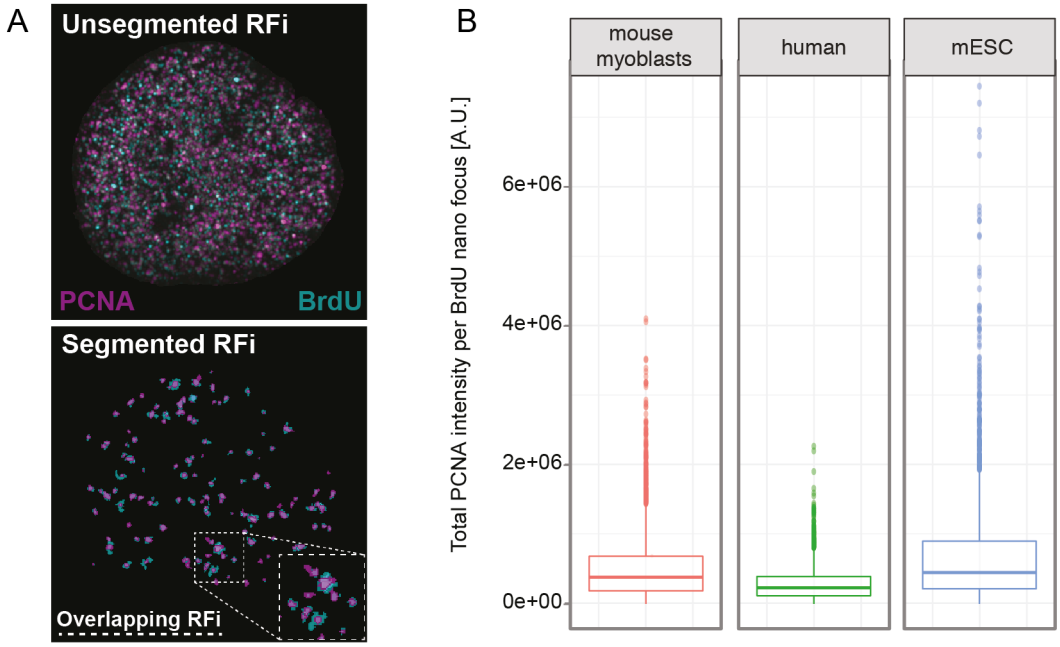


Figure 16 - Are individual replication forks resolved by super-resolution microscopy in ES cells? (A) Maximum projection of 3D-SIM replication signals (upper row; BrdU: turquoise, PCNA: purple). All cell types used in this experiment (mouse myoblasts, human and mES cells) were labeled for 20 min with the nucleotide analogue BrdU and immediately fixed. BrdU and endogenous PCNA were detected by immunofluorescence. Both signals were segmented computationally and overlapping replication signals were used to measure the PCNA intensity within individual BrdU replication signals (lower row, inset in dashed box). The total fluorescence intensity of PCNA was used to measure and compare the number of replication forks in each cell type in order to answer the question if more individual replication forks are resolved in ES cells. (**B**) Boxplots of the PCNA intensities measured within all segmented BrdU replication foci in the different cell lines (N = 3399 (mouse myoblasts), 3231 (human) and 2864 (mESC) overlapping foci measured). Larger dots represent outlier points. The median PCNA intensity is similar between all cell lines (horizontal line within the box, see Tab. S3), indicating that 3D-SIM replication in ES cells represent mostly replicons.

These data, however, suggest that foci of all cell lines contain similar amounts of PCNA. It is possible, however, that this method is not sensitive enough to easily discriminate between one or two replication forks.

Based on the idea that a different chromatin compaction and (loop) organization in ES cells could result in a spatially different organization of replicons, and thus more detectable signals (Fig. 17 B), I aimed to compare the numbers of nanoRFi in a given volume within the cell nucleus of mouse myoblast and embryonic stem cells.

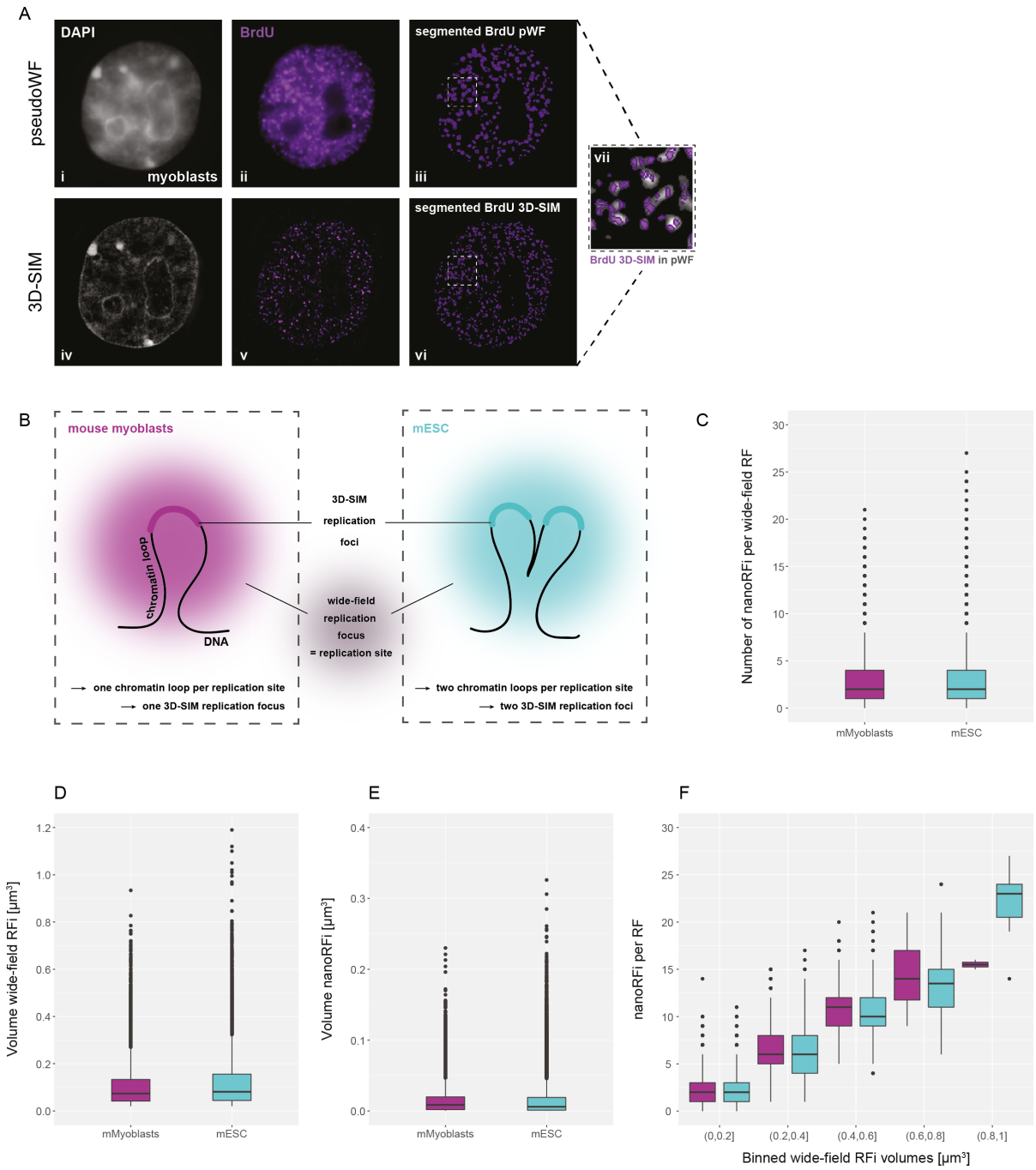


Figure 17 - The number of replicons within a given volume appears to be similar in pluripotent and somatic mouse cells. A different organization at the level of chromatin loops in embryonic stem cells could results in an altered organization of nascent DNA at replication sites. This could allow the detection of more replicons or else individual replication forks by 3D-SIM. (**A**) When imaged at lower optical resolution, replication signals appear as larger foci (pseudo wide-field (pWF) foci). These consist of a number of smaller foci when

resolved by super-resolution microscopy. Shown are representative pWF (upper row) and the respective 3D-SIM images (lower row) of the cell nucleus of a mouse myoblast (left column, **i** and **iv**) as well as the unsegmented (middle column, **ii** and **v**) and segmented (right column, **iii** and **vi**) BrdU replication signals. The pWF replication foci were used to demarcate a distinct volume of DNA in which we quantified the number of nano replication foci (magnified inset, **vii**). (**B**) Graphical outline representing the signals shown in (**A-vii**). pWF replication foci, are represented as blurred signals. Each focus consists of smaller replication sites, i.e. replicons (coloured segments) that are located on the chromatin fiber undergoing replication (black line). If the chromatin loop organization, e.g. into more smaller loops, is different in somatic (mouse myoblasts = purple) and pluripotent cells (mESC = turquoise), we might detect more replicons or replication forks within a given volume, which could explain the higher numbers of measured 3D-SIM signals in mESC. (**C**) Both mouse ES and myoblast cells contain on average three replication nano foci per wide-field focus (mMyoblast: 3.1 ± 0.02; mESC: 3.1 ± 0.2, mean ± SEM). Measurement includes all nanoRFi, irrespective of the volume of the underlying pWF replication foci. A less compact chromatin in ES cells could influence the volume of the pseudoWF foci to which we restricted our measurement. In turn, this could influence the number of nanoRFi measured. Thus, we compared the volumes of both the segmented pWF and 3D-SIM replication foci (boxplots in **D** and **E** and Tabs. S5-S7). No significant difference was observed between the two cell lines. Instead, we observed a linear increase of the number of super-resolved replication signals within increasing volume of the underlying wide-field focus (**F**). Volumes of pWF foci were binned to allow direct comparison between the two cell lines. From this measurement, it seems that the chromatin of somatic and pluripotent cells within a distinct nuclear volume, i.e. the sites at which replication takes place, is organized in the same manner.

As discrete volume(s), I choose segmented (pseudo) wide-field replication foci (pseudoWF-RFi or pWF-RFi) of the same cells. These images can be obtained from the OMX super-resolution imaging system and allow a direct correlative measurement of super-resolved signals in the corresponding (pseudo) wide-field data (see Fig. 17 A).

In earlier studies, wide-field (and later confocal) images were the basis for replication foci number calculations (Chagin et al., 2015; Reinhart and Cardoso, 2017). From these signals it was originally concluded that replication foci correspond to 1 Mbp replication domains that contain clusters of replicons (Berezney et al., 2000). Thus, to directly assess and compare the number of replicons between ES and somatic cells, they represented the most suitable volumetric unit for the intended analysis.

This quantification revealed an almost equal number of nanoRFi within pseudoWF-RFi for mouse myoblast and ESCs (3.1 ± 0.02 for mouse myoblasts and 3.2 ± 0.02 for mouse ESCs, respectively; mean ± SEM; Fig. 17 C). This data contains all nanoRFi measurements, irrespective of the volume of the underlying pseudo wide-field focus. A different chromatin compaction in ES cells, however, could result in a larger volume of the widefield foci, which would influence the measurement of individual 3D-SIM foci negatively.

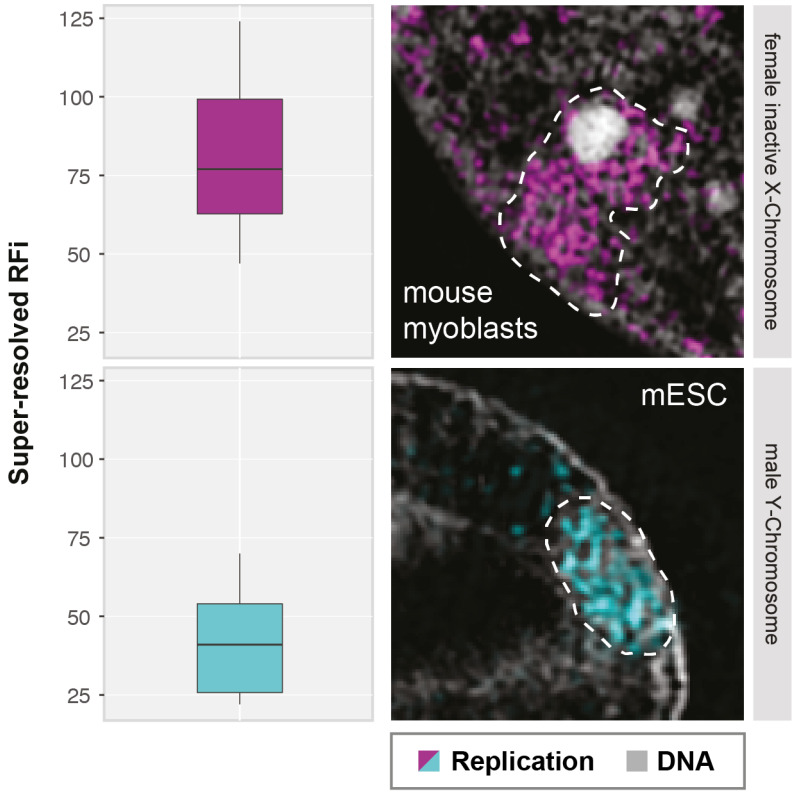
Surprisingly, comparison of both the volumes of both pseudoWF and nano replication foci of the two cell types did not indicate significant differences (see Figs. 17 D and E). Indeed, plotting the number of nanoRFi against increasing volumes of the corresponding wide-field replication foci (different volumes binned equally for both cell types) showed a linear increase (Fig. 17 F), suggesting that the small scale chromatin organization at the level of chromatin loop is, at least in these cells, indeed similar.

Do the female inactive X chromosome and the male Y-chromosome share a common replication mechanism?

The specialized replication timing profile of the *Barr body*, as the Xi is also called, is believed to contribute to the maintenance of the inactive state and with that to the mechanism of dosage compensation in female cells (Koren, 2014; Payer and Lee, 2008)}. The male Y-chromosome has been referred to as a `*functional wasteland*` with only little additional functional relevance for a male organism, besides sex determination during early embryonic development (Quintana-Murci and Fellous, 2001; Quintana-Murci et al., 2001; Sinclair et al., 1990). In that sense, both chromosomes appear to become functionally `shut down` at some point during development.

Interestingly, both chromosomes exhibit similar replication characteristics (synchronous activation of origins marked by a strong accumulation of replication signals) and kinetics (short replication duration of only around one hour, (Casas-Delucchi et al., 2011) and this study). These similarities raise the question of whether the two sex chromosomes share a specific mechanism regulating their mode of replication, maybe as a result of or a requirement of their transcriptionally inactive state.

With the technical possibility to analyze individual replicons along DNA *in situ*, I again focused on the female inactive X and the male Y-chromosome of ES cells to further address the obvious similarities regarding DNA replication. 3D-SIM image data of C2C12 or J1 embryonic stem cells were used to quantify the number of BrdU replication nano foci on the two chromosomes. The results are summarized in Fig. 18.



N =	Mean	Median	SD	SEM	95CI	GS*(Mbp)	$Ratio_{RFi}$ †	$Ratio_{GS}$
6	80.7	77	22.4	5.6	11.9	171.03	1.86	1.87
16	42.2	41	19.2	7.8	20.1	91.74		

* = genome size † = median ratio Xi/Y

Figure 18 - Are the male Y- and the female inactive X-chromosome replicated by a distinct mechanism? 3D-SIM replication foci analysis of female inactive X (Xi; mouse myoblasts; upper row) and male Y-chromosomes (mESC, lower row) revealed further similarities in the mode of replication of the two chromosomes. The number of super-resolved replication foci were compared between the two chromosomes (~81 for Xi- and ~42 for Y-chromosome, respectively; see table). The ratio between these foci counts was almost equal to the ratio of the sizes of the chromosomes (~1.8 in both cases). This finding suggests that the synchronous activation of replication origins on both chromosomes could be related to their respective sizes.

I measured an average number of 40 nanoRFi (42.2 ± 7.8) on the Y- and twice as many on the X-chromosome (80.7 ± 5.6). Due to low sequence mapping coverage, the Y-chromosome is often neglected in genome-wide studies. However, a recent work reported more than 25 000 replication origins that were mapped by sequencing of replication initiation sites (ini-seq) along the human genome, from which only 13 origins were located on the male allosome (Langley et al., 2016), suggesting that the number of replication origins on the Y-chromosome is indeed very low.

Based on the comparison of super-resolved replication foci, it appears that only half the number of replication origins are activated on the male sex chromosome and I wondered if this numbers may be somehow related to the sizes of the chromosomes. Surprisingly, when comparing the ratios of the sizes of X- and Y-chromosome (171.03 Mbp versus 91.74 Mbp, respectively) and nanoRFi, both turned out to be ~1.8 (1.86 in case of replication foci and 1.87 for chromosome sizes derived from the published genome size data used earlier, Tab. 1). This result could indicate the existence of a specialized replication mechanism, for chromosomes that are transcriptionally inactive.

Loss of DNA methylation does not affect the spatio-temporal DNA replication dynamics in embryonic stem cells

The loss of DNA methylation upon knockout or knockdown of the responsible DNA methyltransferases has been shown to result in reactivation of retrotransposable elements (Kimura et al., 2006) and leads to genome instability, growth defects and ultimately cell death in mammalian somatic cells (Jackson-Grusby et al., 2001). Interestingly, pluripotent ES cells are able to tolerate loss of DNA methylation after DNMT1 single- or DNMT3A/DNMT3B double knock-out (Chen et al., 2003; Lei et al., 1996).

It has been shown that hypomethylation, induced by knockout of the *de novo* methyltransferases DNMT3A/B, is accompanied by increased histone acetylation (Jackson et al., 2004), which, in turn, affects chromatin compaction. Similar changes in histone modifications and chromatin architecture upon loss of DNA methylation have been observed in many other studies (Casas-Delucchi et al., 2012; Lehnertz et al., 2003; Tamaru and Selker, 2001; Tariq et al., 2003), and can be explained with changes in common interactions of DNMTs or methylated DNA directly with histone modifiers, like histone methylases (e.g. G9a,) or deacetylases (HDACs) (Esteve et al., 2006; Jones et al., 1998; Nan et al., 1998; Wade et al., 1999; Zhang et al., 1999).

To evaluate the role of DNA methylation as a potential regulator of DNA replication timing in embryonic stem cells, I used a J1-derived stem cell line defective in all three murine DNA methyltransferases (DNMT1/3A/B triple knockout or TKO; (Tsumura et al., 2006)). These ES cells lack global CpG methylation, but were shown to be unaffected in both histone methylation and acetylation marks. Additionally, pericentromeric heterochromatin structure and organization are unaltered, reflected by similar binding of HP1 to and H3K9 trimethylation of chromocenters in wild type and knockout cells. The lack of DNA

methylation in this cell line was confirmed by immunofluorescence staining with an antibody against 5-methylcytosine (Fig. 19).

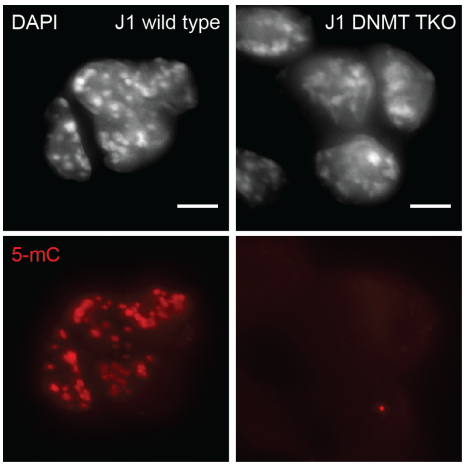


Figure 19 - (**A**) Loss of DNA methylation in DNMT triple-knockout ES cells (J1 DNMT TKO, (Tsumura et al., 2006)) was confirmed by immunofluorescence staining against the cytosine modification 5-methylcytosine (5-mC). In contrast, J1 wt cells showed strong enrichment of 5-mC at pericentromeric heterochromatin. Scale bar = 5 µm.

I also used ES cells deficient for the multi-domain protein Np95, derived from the E14 ES cell line (Muto et al., 2002). Np95 has been shown to preferentially bind hemi-methylated DNA via its SET- and Ring-associated (SRA) domain. Binding of Np95 to heterochromatin is abolished in the DNMT TKO cells that are devoid of DNA methylation (Sharif et al., 2007). Np95 further interacts with all three DNMTs and was shown to be, at least in part, responsible for the recruitment of DNMT1 to replicating pericentromeric heterochromatin. Consequently, $Np95^{-/-}$ cells exhibit similar low levels of CpG methylation as cells lacking DNMT1 (Arita et al., 2008; Meilinger et al., 2009; Papait et al., 2007; Sharif et al., 2007; Uemura et al., 2000). Np95 localizes to replicating pericentromeric heterochromatin in NIH-3T3 cells and knock-down of Np95 in these cells impaired S-phase progression, which was attributed to a role of Np95 specifically in the duplication of heterochromatin.

Knockout and control cell lines were labeled for 15 minutes with EdU and chased for 2 hours with fresh stem cell medium before fixation. EdU and endogenous PCNA were detected as described above. All KO cell lines exhibited the same replication patterns characterized at the beginning of this study. Interestingly, no changes in the temporal order of DNA replication were observe in either of the KO cell lines compared to WT control cells (Fig. 20). This finding is especially interesting with regard to the advanced replication timing of pericentromeric heterochromatin in embryonic stem cells reported in this study.

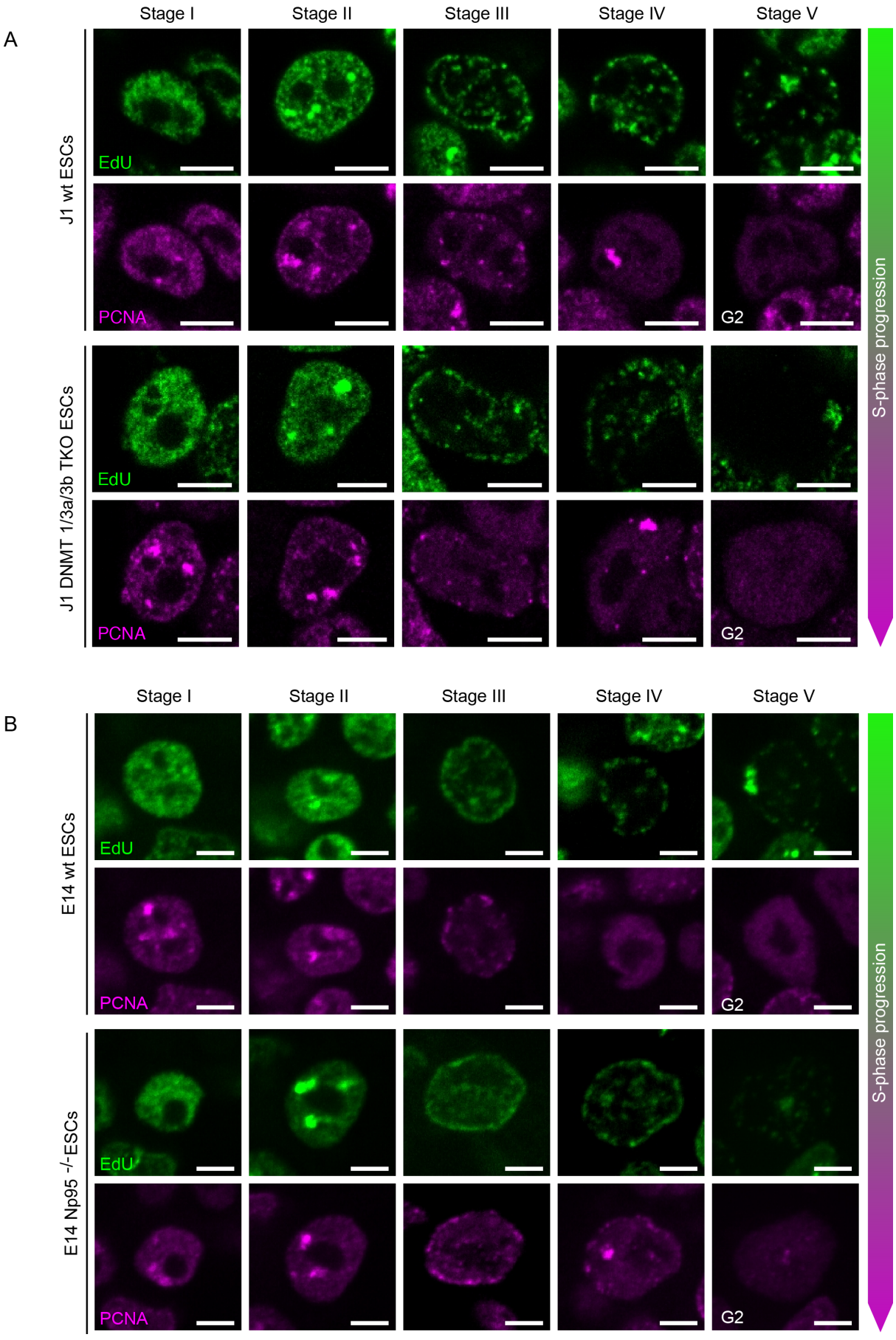


Figure legend on next page

Figure 20 - Loss of global DNA methylation or the key epigenetic regulator Np95 does not affect the spatio-temporal order of DNA replication in ES cells. (**A** & **B**) J1-DNMT TKO, E14-Np95$^{-/-}$ and their respective wild type cells were labeled with EdU (15 min) and fixed after a chase period of 2 hours. Immunofluorescent detection of the nucleotide analogue and endogenous PCNA showed that the loss of DNA methylation did not affect the spatio-temporal order of DNA replication. Progression through S-phase is shown from top to bottom. All replication patterns characterized in the beginning, were observed and their order was unaltered in the respective knockout cells. Scale bar = 5 µm.

To assess, if complete loss of DNA methylation would affect the progression of the cells through S-phase, I then compared the fraction of replicating cells in asynchronous populations of J1 wild type and DNMT triple-knockout cells by determining the percentage of EdU positive cells (Fig. 21A). In both cell lines a comparable fraction of cell were in S-phase (79.3 ± 4.5 % versus 76.2 ± 3.3 % (mean ± SD) in J1 wild type and DNMT TKO, respectively), in line with the results for J1 WT cells shown in Fig. 7. This suggests that S-phase progression in general, is not affected by lack of DNA methylation.

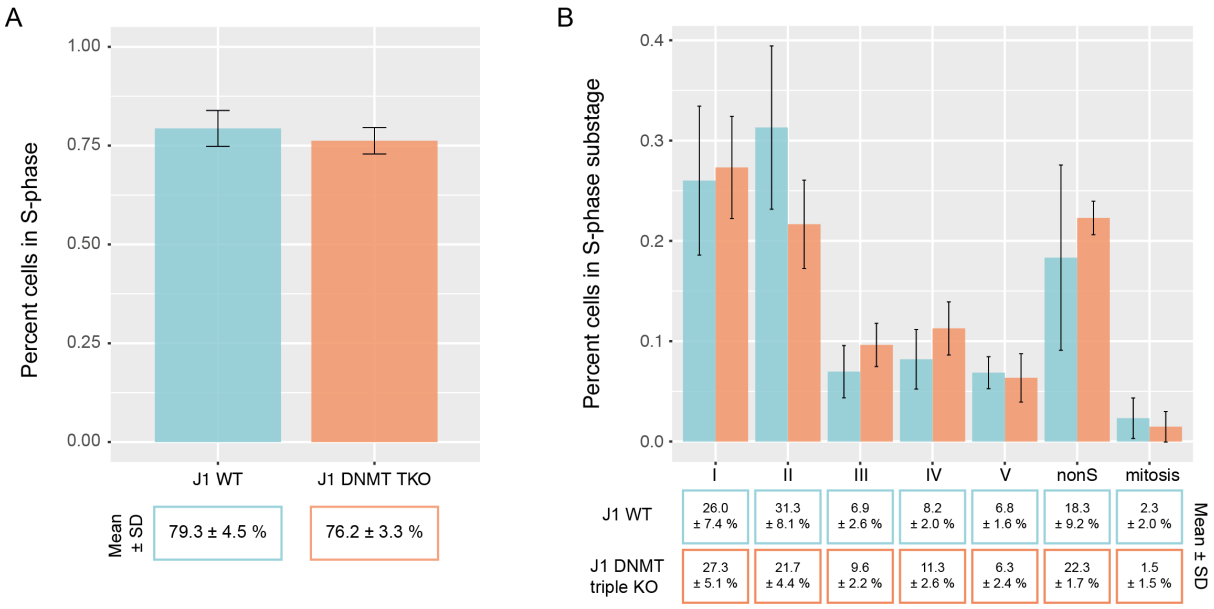


Figure 21 - Loss of DNA methylation leads to faster replication of pericentromeric heterochromatin. (**A**) Bar plots (mean ± standard deviation) showing the percentage of cells in S-phase is similar in J1 wild type and DNMT1/3A/3B triple knockout (TKO) ES cells. The loss of DNA methylation does not significanty affect normal S-phase progression in the TKO cells. (**B**) Analysis of cells at the different cell cycle and S-phase substages (S-phase I - V, G1/G2 (nonS) and mitosis) showed that the fraction of DNMT TKO cells replicating pericentromeric heterochromatin (**II)** is reduced (Bar plots, mean ± standard deviation). This indicates faster progression through the second stage of S-phase upon loss of DNA methylation.

Since pericentromeric heterochromatin contains high levels of DNA methylation in the wild type cells (Fig. 19), I asked if the lack of DNA methylation could affect replication of chromocenters. Therefore, I compared the fraction of cells at all S-phase substages based on the frequency of the different EdU replication patterns (Fig. 21B). I observed a decrease of cells at the second stage of S-phase, indicating that duplication of pericentromeric heterochromatin requires less time than in wild type cells. In turn, the fraction of cells at the later S-phase stages increased accordingly.

Faster S-phase progression can be achieved either by the activation of additional origins of replication or by increasing the replication fork speed. Thus, I compared the replication fork speed of J1 wt and J1 DNMT TKO cells by Molecular Combing. ES cells lacking global DNA methylation (J1 DNMT TKO) exhibit a slight but significant increase in DNA replication fork speed (1.69 versus 1.81 kb/min in WT and DNMT TKO cells, respectively, p-value = 1.83e-5 by Welch Two Sample t-test; Fig. 22).

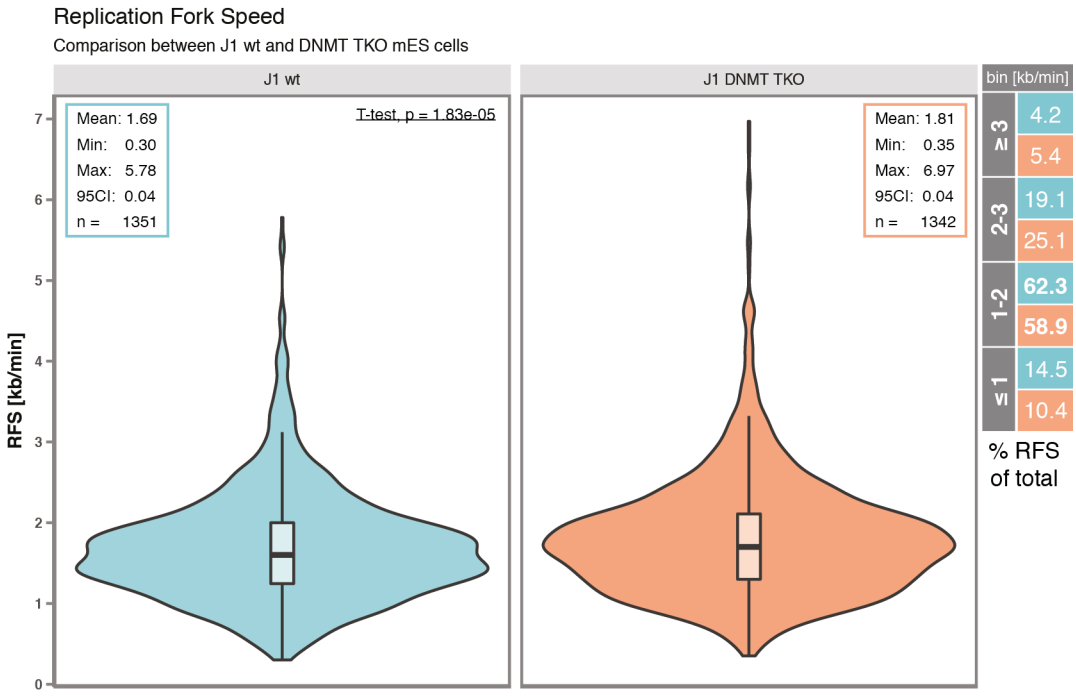


Figure 22 - Loss of DNA methylation advances replication fork speed *in vivo*. Single-molecule measurements in wild type (left) and DNMT triple knockout (right) ES cells revealed a slight but significant increase in replication fork speed upon loss of DNA methylation (mean: 1.69 kb/min (wild type, wt) versus 1.81 kb/min (DNMT TKO), respectively; N = replication tracks measured; p = 1.83e-5 by Welch Two Sample t-test). In both cell lines the measured average for speed represent the largest fraction of all measured RFS (right, % RFS of total).

Together with the observation that DNMT triple-knockout cells require less time for the duplication of pericentromeric heterochromatin, the difference in replication fork speed suggests that the loss of DNA methylation results in faster replication fork progression, specifically along DNA that is usually highly methylated under wild type conditions. However, it remains to be tested whether the lack of DNA methylation also results in increased origin activation within chromocenters.

Developmental differences in the spatio-temporal DNA replication timing program of murine embryonic stem cells

The development and rapid evolution of genome-wide techniques substituted for many classic experimental approaches and also made their way into the field of DNA replication studies. Nowadays, it is possible to derive DNA replication timing profiles (RT-profiles) from large populations of virtually any cell type (Hiratani et al., 2008; Pope et al., 2010). Such profiles revealed that chromosomes consist of distinct *replication domains* with an average size of 1.5 - 2.5 Mb, that are either early or late replicating and that exhibit sharp boundaries between neighboring domains with different replication timing (Hiratani et al., 2008). These datasets can be directly linked to the underlying DNA sequence, which makes genome-wide experiments powerful in the sense that they provide a useful link with information about cell type specific gene expression profiles or data on DNA methylation and histone modifications, respectively.

It was recently shown that the chromosomes of embryonic stem cells consist of a large number of small replication domains (Ryba et al., 2011) and that up to ~50% of the pluripotent mouse genome experiences changes in its replication timing during differentiation (Hiratani et al., 2010; Hiratani et al., 2008). Specific RT-profiles were used to discriminate different cell types, and, together with the observed large scale rearrangements upon lineage commitment of differentiating ESCs, it was suggested that genome-wide RT-profiles may serve as a novel fingerprint for cellular identity and even pluripotency (Ryba et al., 2011).

However, such experiments divide cell populations into either early or late replicating fractions which results in a broad population average. On the other side of the coin, single-cell methods allow the analysis of RT-profiles of only a handful of genomic loci (Van der Aa et al., 2013). Either way, they exhibit a relatively poor temporal resolution with respect to the underlying kinetics of molecular processes and are incapable of capturing additional important parameters at the level of individual cells. Furthermore, any structural and organizational information of the underlying chromatin is lost as the data is derived directly from immunoprecipitated DNA of FACS sorted cells, although the combination with data from chromosome conformation capture experiments revealed that replication timing domains show good correlation with either A- or B-compartments (Pope et al., 2014; Ryba et al., 2010).

With the pulse-chase experiments performed in this study, I was able to contribute to a more complete understanding of the DNA replication dynamics in embryonic stem cells, thereby supplementing genome-wide data. I first presented the identification and characterization of a total of five replication patterns that specifically mark different substages of S-phase and that correspond to the duplication of different chromatin types as well as an entire chromosome in pluripotent cells (Fig. 4 and 5).

As in somatic cells, euchromatin is replicated at the beginning of S-phase as seen by the homogeneous distribution of replicating sites in the cell nucleus (Fig. 5, top row). At this early stage, only peripheral chromatin regions as well as clusters of pericentromeric heterochromatin are devoid of any replication label.

Surprisingly, I found that pericentromeric heterochromatin replicates much earlier in murine embryonic stem cells than in somatic cells. In the pulse labeling experiments, this became apparent by the accumulation of replication signals at chromocenters in DAPI-stained interphase nuclei. I later confirmed this result with the use of major satellite specific FISH probes in a separate Repli-FISH experiment (Fig. 10 A). Heterochromatin usually replicates late during S-phase in most organisms studied. However, also exceptions to this rule have been found (Ahmad and Henikoff, 2001; Kim and Huberman, 2001; McCarroll and Fangman, 1988; Vig, 1995). In *Drosophila*, for example, it was shown that satellite sequences became increasingly heterochromatic and late replicating only with successive differentiation at later developmental cycles. Interestingly, this late replication even seemed to be responsible for the prolongation of S-phase which ultimately allowed the progression to the midblastula transition accompanied by the characteristic changes in cell cycle organization upon differentiation (Shermoen et al., 2010).

A different chromatin organization, together with epigenetic modifications of histones, may results in less compact chromatin, which could fine-tune molecular processes by providing better access for the required protein machineries to reach their template (Gaspar-Maia et al., 2009; Hargreaves and Crabtree, 2011). As a results, this could increase the overall efficiency, speed or timing of such a process.

Chromatin in embryonic stem cells is generally less compact and chromatin associated proteins are highly mobile in ES cells (Lopes Novo and Rugg-Gunn, 2016; Meshorer and Misteli, 2006). Similar observations have been described in previous studies ((Efroni et al., 2008; Meshorer et al., 2006)) and for differentiated neuronal precursor cells, in which pericentromeric heterochromatin was visualized with probes specific for the major satellite

repeats (Brero et al., 2005). It is considered that such open conformation represents a necessary prerequisite of pluripotent cells to remain responsive to the changes that occur during differentiation (Gaspar-Maia et al., 2009). The results presented here support the finding that, especially pericentromeric heterochromatin, is decompacted in ES cells as reflected by their diffuse structure in the stem cell nucleus (Fig. 11).

Different layers of epigenetic modifications exist, including the association with heterochromatin-specific proteins (e.g. HP1) and increased DNA methylation, that are responsible for constitutive heterochromatin formation and maintenance. How exactly the replication machinery penetrates this tightly compacted chromatin structure is not fully understood. It has been proposed that the duplication of pericentromeric heterochromatin is mediated by a specific unit termed the *pericentromeric heterochromatin duplication body* (pHDB, (Quivy et al., 2004)). This unit was shown to consist of PCNA, the chromatin assembly factor CAF1 as well as a specific *heterochromatin binding protein 1* subpopulation (HP1α), all of which associated with the periphery of mouse pericentromeric heterochromatin during S-phase. Based on nucleotide labeling experiments, it was suggested that major satellite DNA that undergoes replication, gets located to the periphery of a chromocenter, looped out from its core and is incorporated back into the heterochromatic clusters after duplication by the pHDB. In support of a role for chromatin remodelling complexes in heterochromatin replication, *Collins et al.* showed that the chromatin remodeling complex ACF1-ISWI supports better access for the replisome through decondensation of heterochromatin and that DNA replication progression in late S-phase was impaired upon depletion of the complex in mouse cells (Collins et al., 2002). Given that different chromatin remodeling complexes are expressed at higher levels in ESCs (Kurisaki et al., 2005) it would be interesting to investigate their exact role in the regulation of heterochromatin compaction and, as a consequence, the efficiency of DNA replication of this heterochromatic compartment. Surprisingly, treatment of cells lacking ACF1-ISWI with the demethylating agent 5-azacytidine (5-aza) fully reversed the replication defects. Changing the levels of DNA methylation with 5-aza was previously shown to affect heterochromatin condensation (Haaf and Schmid, 2000). This indicates that, indeed, the increased compaction of heterochromatin upon loss of the remodelling complex was responsible for the accumulation of cells in late S-phase.

The binding of specific readers of methylated cytosines (5-methylcytosine, 5mC) is known to influence gene expression and also chromatin structure (Clouaire and Stancheva, 2008). J1 ES cells were shown to contain higher levels of 5-hydroxymethylcytosine (5hmC) due to the action of ten-eleven translocation enzymes 1 and 2 (Tet1 and Tet 2) that convert 5mC

to 5hmC and which are both expressed in mESCs (Dawlaty et al., 2013). A recently published study further uncovered a role of the methylcytosine-binding protein 1 (MBD1), which is also expressed in embryonic stem cells (Kobayakawa et al., 2007), in the process of demethylation of 5mC (Zhang et al., 2017b). Overexpression of MBD1 was shown to recruit Tet1 to clusters of pericentromeric heterochromatin, which increased the levels of 5-hydroxymethylcytosine. This argues for a potential role of MBD1 and Tet1 in the maintenance of a decompacted heterochromatin in mESCs, which, in turn, could decrease the binding of 5mC readers and lead to a decrease in compaction.

As hypoacetylation of the core histones H3 and H4 are usually a characteristic feature of heterochromatin and crucial for its maintenance as well as that of silenced genes (Almouzni and Probst, 2011), acetylation of histone tails are another, very potent candidate for the regulation of chromatin compaction and replication timing. Previous experiments from our lab, indeed showed that the replication timing of chromocenters can be advanced by increasing the global level of histone acetylation. Importantly, this effect was independent from DNA or histone methylation (Casas-Delucchi et al., 2012). After drug induced histone hyperacetylation (TSA treatment), the authors observed a direct decompaction of the heterochromatic clusters and concluded that this led to the advanced DNA replication timing. Notably, an earlier replication timing was also observed in cells that were hypomethylated due to loss of the methyltransferase DNMT1 (Casas-Delucchi et al., 2012) but this effect could be attributed to a simultaneous increase in histone acetylation at chromocenters.

In addition to the idea that differences in chromatin mediate an earlier replication timing, a mechanistic basis comes from multiple studies that reported that the *origin recognition complex* (ORC) interacts with the repressive histone marks H3K9m3, H3K27m3 and H4K20m3, which are enriched in heterochromatin (Bartke et al., 2010; Pak et al., 1997; Vermeulen et al., 2010) and that interaction of ORC with HP1 might even be required for heterochromatin formation (Prasanth et al., 2010). These data suggest that the ORC might enrich specifically at heterochromatic regions during origin licensing. This, although counterintuitive, could represent a suitable mechanism for cells, to allow sufficient licensing of chromatin that is usually less accessible, providing an interesting link between the degree of chromatin compaction and the advanced replication timing of pericentromeric heterochromatin seen in ES cells. If mESC exhibit a similar binding of ORC, then the combination with heterochromatin decompaction would allow better access for origin activation factors and could lead to an earlier replication of chromatin enriched for origin complexes.

In this context, it would be interesting to investigate the biological relevance of an earlier replication timing of heterochromatin in ES cells. With regard to their special cell cycle characteristics (i.e. short G1-phase), this could represent some sort of safety mechanism to ensure sufficient duplication of an otherwise hard-to-access chromatin compartment, ahead of the end of S-phase.

In summary, decompaction of heterochromatin is likely to have strong effects on its DNA replication timing. In this, easier access and binding of replication licensing and activation factors may facilitate easier progression along chromatin during S-phase. Therefore, such differences, mediated by increased histone acetylation levels and/or an increased availability or activity of chromatin remodeling complexes in embryonic stem cells, could allow for replication of pericentromeric heterochromatin in the first half of S-phase.

Replication of facultative heterochromatin, which is located at the nuclear periphery, takes place during mid S-phase in somatic cells. Previous studies have shown that this localization is highly conserved between ESCs, differentiated neuronal precursor cells (NPCs) and terminally differentiated astrocytes (ACs, (Wijchers et al., 2015)). This suggests that there is no major reorganization of this inactive compartment during lineage commitment and I therefore concluded that, similar to differentiated cells, facultative heterochromatin in embryonic stem cell gets replicated in the middle of S-phase (Fig.5 - III).

The replication signal distribution that followed duplication of facultative heterochromatin (Fig. 5 - IV) was unrelated to any pattern we had seen in other cell types before. Based on their morphology and perinuclear distribution, however, I assumed that these signals would correspond to some sort of heterochromatin, i.e. centromeric or telomeric DNA. Surprisingly, none of the repeat specific probes used in *fluorescence in situ hybridization* experiments overlapped with these replication sites (Fig. 10 - B & C).

Hozak et al., described similar signal aggregates toward the end of S-phase in cells analyzed by electron microscopy and speculated that they might represent sites were residual unreplicated DNA could be processed before S-phase is completed (Hozak et al., 1994). This remains an attractive explanation, as high amounts of unreplicated DNA could be a result of the short gap phases and frequent cell divisions of embryonic stem cells. On the other hand, it has recently been reported that mouse embryonic stem cells fail to accumulate 53BP1 nuclear bodies (NBs) due to a short G1-phase (Ahuja et al., 2016). In somatic cells these NBs form, in order to sense residual damage sites resulting from

replication stress or under-replicated DNA. Localization of DNA to this specialized repair sites allows their sufficient repair before or after a new round of genome duplication (Lukas et al., 2011). In light of these findings, it would be exciting to investigate if the replication signals observed during the last part of S-phase in mESCs belong to residual unreplicated DNA and how the cells make sure that they will not lead to the loss of important genetic information over the next cell cycle(s).

Late replication timing of the male sex chromosome and similarities to the inactive X-chromosome of female cells

The mechanism of X-chromosome inactivation during the early development of female cells is a process that attracts a lot of interest, especially for the study of epigenetic silencing mechanisms and has been investigated extensively (Cheng and Disteche, 2004; Deng et al., 2014; Okamoto et al., 2004; Takagi and Sasaki, 1975). Inactivation of one of the X homologues in female mammals is required to compensate for gene dosage differences that would otherwise exist over male cells, which carry only one X-chromosome. It was previously shown that the inactive homologue of the female sex chromosome replicates synchronously and within 1-2 hours during early mid-S-phase in mouse myoblast cells. This replication timing is different from the active X-chromosome homologue in the same cell (Casas-Delucchi et al., 2011). How a synchronous activation of licensed replication origins on only the Xi is controlled is currently unknown, but it was shown that some origins are shared on both the active and inactive homologue (Rowntree and Lee, 2006), which suggests that epigenetic mechanisms could play an important role for the synchrony of origin firing. The fact that the inactive X-chromosome accumulates several of these marks, such as increased methylation of histone 3 lysine 27 (H3K27m3), histone hypoacetylation and the binding of the non-coding (nc)RNA *Xist* (Casas-Delucchi et al., 2011), represents a profound basis for this assumption. However, although X-inactivation is observed in many mammalian species like mice, humans and marsupials, they differ in the epigenetic modifications that are associated with the inactive state and also in the mechanism by which inactivation is achieved (Sado and Sakaguchi, 2013). Thus, their direct role in the coordination of synchronous firing of origins remains unclear.

The Y-chromosome of male cells, one of the smallest chromosomes in human and mice (~57 Mbp and ~92 Mbp, respectively; for mouse see Tab. 1), is usually only studied in the context of evolution, clinics and forensics (Jobling and Tyler-Smith, 2003), but otherwise believed to consist of non-functional pieces of DNA (junk DNA; (Singh et al., 2011)). Compared to all other chromosomes it contains the lowest number of genes (Aken et al.,

2016) of which only a fraction appears to be potentially protein coding in human and that are mostly required for testis development and sex determination. Other than that, the Y-chromosome is mostly heterochromatic and is frequently lost in most male cell lines during prolonged cell culture. As a consequence of its nondescript existence, it is often neglected in research studies.

In this work, it was shown that the male sex chromosome marks the end of S-phase in differentiated and pluripotent male cell types (MEFs and embryonic stem cells) as shown by Y-chromosome specific 3D-FISH with simultaneous replication staining (Repli-FISH). This mode of replication, at least in the ES cells, occurs in an equally synchronous manner as that of the Xi and in a short time frame of more or less 1 hour (see Figs. 6 & 7).

Analyzing and comparing the number of replication nano foci on the embryonic Y- and female inactive X-chromosome of mouse myoblast cells by 3D structured illumination microscopy uncovered further similarities. The number of replication sites on each chromosome correlated well with their respective size, suggesting the existence of a common and special mechanism of replication for the two chromosomes. The licensed pre-replication complexes that are loaded onto chromatin at the end of mitosis and that mature as DNA replication proceeds, provide the molecular basis for such mechanism, which might then become synchronously activated at the particular times during S-phase.

Both chromosomes are found in association with the nuclear periphery. This, at least in the case of the inactive X-chromosome in *Mus musculus*, is a results of the high abundance of constitutive lamina-associated domains (cLADs) scattered within its DNA sequence (Solovei et al., 2016). These sequences form the basis for interactions with the nuclear lamina, specifically the Lamin B-receptor (LBR) and the binding to the nuclear periphery is known to be involved in the regulation of (late) heterochromatin replication. This, however, does not provide an explanation for the synchrony of origin activation *per se*.

Notably, such synchronous replication is in stark contrast to the suggested domino-like activation mechanism by which DNA replication is otherwise thought to propagate throughout the genome. Factors that might be responsible for this tightly organized activation are unknown, but could involve the absence of transcriptional activity, resembling the uncontrolled firing events in *Xenopus* and *Drosophila* embryos before the onset of transcription at the mid-blastula transition. Such a relation has been suggested earlier for the inactive X-chromosome (Casas-Delucchi et al., 2011), and could be explained by the absence of topologically associating domains (TADs) on the Xi (Deng et al., 2015), which

are important chromosomal features for the regulation of transcription. In this respect, it would be interesting to more specifically analyse and compare the synchrony of replication of the Y-chromosome in differentiated cells. If gene expression is a driving force for the simultaneous activation of replication origins, differences in the Y-chromosome expression profile between developmental stages could also influence this synchrony, resulting in desynchronized replication of at least some regions of the Y-chromosome, again similar to the differences seen in case of the active and inactive homologue of the X-chromosome in female cells (Casas-Delucchi et al., 2011).

Two recent studies that performed a genome-wide mapping of replication origins with different methods, consistently reported that the number of mapped origins did not correlate with the length of the chromosome (Besnard et al., 2012; Langley et al., 2016). Thus, with regard to the relation of replication nano foci and the sizes of the Xi- and Y-chromosome observed in this study, it would be interesting to investigate whether the synchronous activation itself could limit the amount of required origins. Since all origins are activated within a distinct nuclear region, the recruitment and accumulation of the required biochemical factors would allow efficient activation and replication, representing a *sped up* and efficient version of the domino-like propagation. If, indeed, changes in the synchrony of replication of the Y-chromosome exist upon lineage commitment this question could be addressed for example by comparing the number of nanoRFi in both ES and differentiated cells.

Interestingly, the pericentromeric regions of both chromosomes appears to be uncoupled from this synchronous replication mechanism as they appeared to replicate together with those of the remaining chromosome in male mES, murine embryonic fibroblast cells and female mouse myoblast cells (Casas-Delucchi et al., 2011). This strongly argues for independent mechanisms that control the duplication of pericentromeric heterochromatin and chromosomes that undergo heterochromatization at some point during development, respectively. How this deviating replication timing can be achieved is not clear but could involve heterochromatin binding protein 1. HP1 is conserved from yeast to humans (Eissenberg and Elgin, 2000), specifically recognizes and binds the trimethylated lysine of histone H3 (H3K9m3, (Lachner et al., 2001)) and is involved in heterochromatin establishment and maintenance. However, it has also been shown to advance the replication timing of pericentromeric heterochromatin and the silent mating-type locus of *S. cerevisiae* via direct interactions with the kinase DDK that is responsible for origin activation (Hayashi et al., 2009), and pericentromeric heterochromatin of *D. melanogaster* replicated earlier upon loss of HP1 after siRNA-mediated knock-down (Schwaiger et al., 2010). Thus,

if replication of (peri)centromeric heterochromatin is indeed controlled separately as suggested above and if HP1 binding to chromocenters on the inactive X- and Y-chromosomes is similar to other chromosomes, it could be a suitable candidate to explain the observed replication timing differences.

The maintenance of the inactive X-chromosome depends on the expression and binding of *Xist* RNA from the X-inactivation center, histone hypoacetylation, high H3K27 trimethylation and a high degree of condensation (Casas-Delucchi and Cardoso, 2011). It would be interesting to see if the Y-chromosome also exhibits an epigenetic landscape which could provide better functional insight of this special piece of DNA. *Singh et al.* published data on a few post-translational histone modification found in the euchromatic portion of the human Y-chromosome as well as information on binding sites for CTCF within these sequences (Singh et al., 2011). CTCF plays an important role in the context of chromatin (loop) organization. Although we did not yet specifically investigate histone modifications and a potential effect on the replication timing, we propose that both the Y- and the inactive X-chromosome in female cells exhibit a specialized mode of replication that deviates strictly from the idea of a sequential activation of origins over the course of S-phase. Such mechanism might be important, e.g. to maintain the heterochromatic state of both chromosomes or other, yet undiscovered functions.

Given its distinct peripheral localization in the cell nucleus, the special replication timing and more detailed knowledge about binding sites of factors like CTCF and other important chromatin organizers in the future, the Y-chromosome might evolve to an interesting model to address DNA replication in the context of chromatin organization. Furthermore, we might be able to derive new evidence for the basis of mammalian replication origins by mapping these sites within these two, truly special `chromosome territories`.

Replication timing of centromeres and telomeres in murine ES cells

Although the compaction and dynamics of chromatin in the nucleus of cells at different developmental stages can differ substantially, the global organization of individual chromosomes is by no means random. Instead, they are packed and distributed in a manner that results in the functional separation of eu- and heterochromatin into distinct compartments (A- and B-compartments; (Solovei et al., 2016)). Also, multiple subnuclear compartments exist that form at specific locations in the cell nucleus to concentrate different factors, in order to perform specialized functions like mRNA maturation in splicing speckles

or the transcription of rDNA that takes place in the nucleolus (Cardoso et al., 2012). Repetitive DNA sequences, which make up half of the human genome and around 40% of the mouse genome, also play important roles, e.g. during sister chromatid segregation (centromeres) or to protect the chromosomal ends (telomeres), as described at the beginning of this work.

Mixed results have been obtained for the replication timing of telomeres and centromeres, although most heterochromatin is generally found late replicating that might be linked to nuclear position and which is to be important for its correct function during the cell cycle and (Csink and Henikoff, 1998; Heinz et al., 2018). The centromeres of *Drosophila* as well as yeast, on the other hand, were found to replicate only early in S-phase and this observations suggest that early replication timing is a conserved feature of centromeres (Ahmad and Henikoff, 2001; McCarroll and Fangman, 1988). Weidtkamp-Peters *et al.,* however, published that replication of minor satellite repeats occurs throughout S-phase in mouse fibroblast cells (Weidtkamp-Peters et al., 2006). (Kim et al., 2003; Kim and Huberman, 2001)

Centromeric DNA of ES cells was found to also replicate preferentially early in S-phase (Stage II), together with major satellite repeats. Given their physical location along chromosomes and their 3-dimensional organization in interphase nuclei (Fig. 9A & D and (Guenatri et al., 2004)), it is likely that centromeric regions are also less compacted in ES cells so that partial overlap of the replication timing profiles of these two structures can be expected. A study performed in yeast, found that inner centromeric sequences are early replicating (Kim et al., 2003; Kim and Huberman, 2001) and are able to influence the activation of late firing origins towards an earlier time point when placed in their proximity, suggesting that centromeres themselves could play a role in the regulation of origin activation (Pohl et al., 2012). Given their importance in chromatid segregation during mitosis, early replication of centromeric and pericentromeric regions, might be a necessary property. In embryonic stem cells that possess a cell cycle organization that deviates highly from that of somatic cells, this feature would be even more important in order to ensure full and correct duplication of these hard to replicate structures.

Telomeres fulfill an enormously critical role in the protection of every chromosome by forming complex G-quadruplex structures composed of G-rich repeat DNA that associate with specialized proteins like *shelterins* and are involved in cellular ageing and senescence (Blackburn, 2001). During each round of replication, however, these complex structures are difficult to duplicate by the replication machinery and even require specialized helicases like

Blm (Sfeir et al., 2009) or *Wrn* (Crabbe et al., 2004). As a consequence, a substantial fraction of telomeric DNA is lost due to insufficient duplication of the repeats. Although yeast telomeres are replicated together at the end of S-phase, multiple studies could show that mammalian telomeres replicate throughout S-phase ((Arnoult et al., 2010; Wright et al., 1999; Zou et al., 2004)).

The telomeres of murine embryonic stem cells were also found to replicate over the whole course of S-phase (Fig. 10 C). Although we did not further investigate how this broad replication timing could be regulated, the time of replication has been related to nuclear position in earlier studies, with telomeres positioned towards the nuclear interior replicating earlier than the once associated with the nuclear periphery (Arnoult et al., 2010). In yeast, unlike the usually late replicating chromosome ends, short telomeres get replicated early in S-phase which provide sufficient time for the telomerase enzyme to achieve proper extension (Bianchi and Shore, 2007) in order to prevent further telomere shortening which would otherwise have serious consequences for the cell. It could be possible that differences in telomere length also influences their replication timing in embryonic stem cells that also express the telomerase enzyme (Hiyama and Hiyama, 2007), leading to the observed replication timing distribution.

Analysis of replicons in murine embryonic stem cells may suggest further developmental differences in DNA replication dynamics

Replicons are the smallest functional unit of DNA replication and are found to be organized as clusters along the DNA fiber that share a similar replication timing (Jackson and Pombo, 1998). *In situ*, clusters containing on average 2-10 replicons, become visible as discrete replication foci, representing domains of ~1 Mb of DNA (Berezney et al., 2000).

For a long time, the replication of these clusters, or tandems thereof, was believed to take place at distinct sites within the cell nucleus, where aggregates of replication factors attached to an underlying nuclear scaffold form `replication factories` (Hozak et al., 1994).

In contrast to this *replication factory model,* Chagin and colleagues could recently show that replication foci (RFi) in somatic cells of mice and humans correspond to individual replicons, rather than static accumulations of replication factors (Chagin et al., 2016), and concluded that each nano replication focus thus represent the smallest resolvable functional unit in the context of DNA replication. Hereby, the close association of replication

signals at distinct sites is thought to depend on the underlying organization of the chromatin fiber, but not the trapping of replication factories from which the replicated DNA then extrudes.

Chromatin loop structures are themselves considered distinct units in the complex hierarchy of organization within the cell nucleus. At this level, individual loops may harbor the DNA elements that act as templates or regulators in different molecular processes. The most prominent examples are enhancer elements and promoters of genes that coordinate and regulate transcription via long ranging *cis*-interactions (Fraser, 2006; Krivega and Dean, 2012; Tolhuis et al., 2002). Thus, differences in loop conformation and (local) chromatin density are potent features that can modulate interactions between genetic elements or regulate the accessibility of the biochemical factors that are required for DNA-dependent metabolisms and, consequently, the processes as a whole.

The changes in chromatin organization and the accumulation of condensed heterochromatin observed upon lineage commitment are accompanied with the establishment of a cell type specific gene expression program (Hiratani et al., 2008). With these changes, also the replication timing of pluripotency related genes, like OCT4 and NANOG, correlates well with their transcriptional activity in ESCs and switches to late replication upon differentiation, while cell type specific genes of differentiated cells shift to early replication (Perry et al., 2004).

Similar to the regulation of transcription, frequent topological interactions of DNA in *cis* or even *trans* could be able to also regulate DNA replication. Indeed, it was shown that topologically associating domains (TADs) are themselves regulatory units of DNA replication timing (Pope et al., 2014), which is in agreement with idea of a domino-like activation of replication origins. Hereby, activation of inactive origins is achieved by the transient spatial association with already activated origins that create a local decondensed chromatin environment, permissive for further activation (Casas-Delucchi and Cardoso, 2011; Chagin et al., 2010).

In this study, I aimed to evaluate whether replicons and their organization are conserved between pluripotent and representative somatic cells that were used in a recently published study (Chagin et al., 2016). I performed a detailed analysis and characterization of the cell cycle characteristics of pluripotent embryonic stem cells, which are known to exhibit many differences to lineage committed cell types (Ballabeni et al., 2011).

Most cells in an asynchronous population of the embryonic stem cell line used in these experiments were frequently undergoing DNA replication, reflected by the high fraction of cells that stained positive for nucleotide analogues incorporated into nascent DNA (Fig. 7C). Analysis of the population doubling time allowed to calculate the length of S-phase as well as the individual cell cycle and S-phase (sub)stages, and showed that the cells exhibited a short G1-phase, while the S-phase duration was comparable to the frequently reported 8 - 10 hours for somatic cells (Chagin et al., 2016; Hahn et al., 2009).

The analysis of the molecular properties of stem cell replicons revealed that the processivity of replication forks of ES cells is comparable to human cancer cells. The distance between neighboring replication origins (inter origin distance, IOD), however, was smaller than in the somatic mouse and human cells, but within the range of observed inter origin distances in other cell lines (on average 50-150 kb, (Berezney et al., 2000)). The inter origin distance is frequently used to indirectly estimate the density of activated origins in a cell and a shorter IOD can be interpreted as more frequent origin activation events, thus more activated origins in total, considering a homogeneous distribution or replicons or replicon clusters, respectively (Conti et al., 2007; Maya-Mendoza et al., 2007; Techer et al., 2013).

Integration of these parameters allowed the calculation of the total number of replicons activated within the approx. 9 hours of S-phase in mESC as well as the average number of replicons active in parallel at any time during S-phase (Tab. 3 & 4). In combination with the results from the 3D-SIM replication foci measurements, two interesting conclusion were obtained. i) embryonic stem cells activate a comparable amount of replication origins over the course of S-phase as the somatic cells with an substantially larger genome size (Tab. 3) and ii) in contrast to somatic cells, in which one super-resolved replication focus corresponds to approximately one replicon (Tab. 4), the ratio between simultaneously active replicons and nanoRFi was significantly smaller in mESCs.

The total number of replicons required for full genome duplication can be estimated by the ratio of genome size of the diploid ES cell line and the average IOD, which can be used as an approximation for replicon size. Based on the inter origin distance measurements, embryonic stem cells activate around 57,714 origins of replication during S-phase, a number that is comparable to that of the polyploid somatic cell lines (51,404 and 70,501 for human and mouse, respectively, see Tab. 3 and (Chagin et al., 2016) for reference). As replication fork speed and/or the number and spacing of origins determine the time required for genome duplication, i.e. the duration of S-phase, the values obtained by *Chagin et al.* could be specific to the cell lines used and may be a result of their polyploid karyotype.

Indeed, the authors concluded that the C2C12 mouse myoblast cell line compensated for differences in genome size by increasing the rate of fork movement rather than the activation of additional origins. In turn, this could mean that a somatic cell line with a similar DNA content than the ES cell line used in this study, could either exhibit differences in replication fork speed or in the number of activated origins, and with that nano replication foci, or both.

It has been shown multiple times that a higher rate of active origins of replication are formed on the chromosomes of embryonic cells than on that of somatic cells from other species. The best example, again, is the replication of DNA within *Xenopus* egg extracts or *Drosophila melanogaster* embryos that occurs very efficiently from a large number of origins that are on average only 10-15 kb apart. (Blumenthal et al., 1974; Callan, 1974). Several studies have been published that provide data that could explain the hypothesis that also murine ES cells activate more replication origins and they represent the basis for future experiments.

The first step that represents the foundation for potential differences in origin firing is the licensing of DNA that occurs at late M/G1-phase of the cell cycle and that begins with the loading of ORC complexes onto potential replication origins (Dimitrova and Gilbert, 1999). This is followed by the subsequent recruitment of CDC6 and the CDT1-bound MCM2-7 helicase complexes to form the pre-replication complex. As mammalian replication origins are not solely defined by an underlying sequence motif and given that ORC does not exhibit definite sequence specificity, its association with DNA, and with that the spacing between pre-RCs, is dynamic. Although licensing of origins occurs in excess, only a subset will be activated over the course of S-phase by the coordinated action of the S-phase specific kinases CDK and DDK (Bousset and Diffley, 1998; Dowell et al., 1994). Thus, the choice about origin density can be controlled at multiple levels, which means that even when the number and spacing of pre-replication complexes along DNA would be the same in embryonic and differentiated cells, the number and sizes of the resulting replicons will ultimately depend on the number of pre-RCs that will be activated during S-phase. This idea is illustrated in Fig. 23.

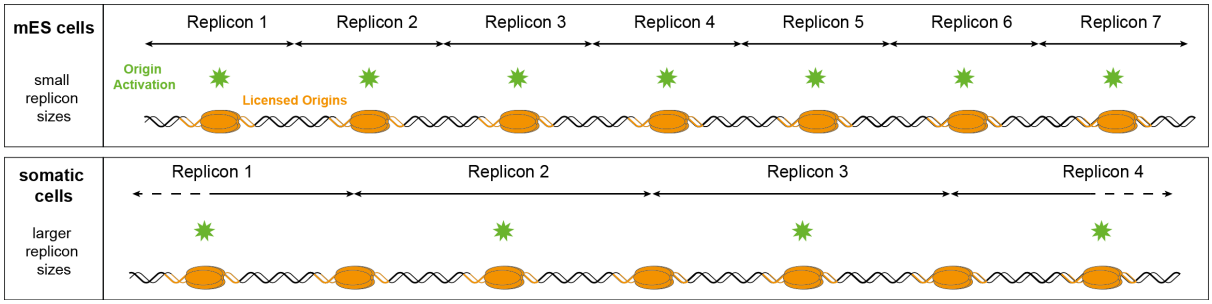


Figure 23 - How an increased origin activation might explain the discrepancy between replicons and nanoRFi in ES cells Experiments comparing developmental stages of *Xenopus laevis* revealed differences in the activation and spacing of origins of replication during S-phase. Early embryonic stages initiate DNA replication from numerous sites that are only 10-15 kb on average apart. Upon differentiation, the number of activated origins decreases. We propose a similar mechanism for mouse embryonic stem cells. Along a given segment of chromosomal DNA carrying multiple licensed origins, pluripotent stem cells might initiate DNA replication from more sites, e.g. all seven origins shown in this drawing (upper row) than somatic cells (lower row), which might fire only every second origin. This would result in more, but smaller individual replicons and shorter inter-origin distances, e.g. as measured in our experiments for mESC (~90 kb) and in (Chagin et al., 2016) for differentiated mouse and human cells (161.7 kb (C2C12) and 188.7 kb (HeLa), respectively). More replicons would also lead to higher numbers of super-resolved replication foci in embryonic stem cells as seen in our experiments. Furthermore, this could explain the discrepancy in the relation of 3D-SIM RFi and replicons that we obtain for stem cells and that supposedly contradict the interpretation from *Chagin et al.*

In an elegant study, *Walter et al.* originally addressed the regulation of the developmental changes in replicon size in *Xenopus*, and showed that the loading of ORC complexes remained constant in extracts containing increasing numbers of nuclei/µl, which mimicked the increase of nuclei within the developing embryo (Walter and Newport, 1997). Instead, the authors found that the number of initiation events decreased in reactions containing concentrations of 2000 or more nuclei/µl, which lead to an significant increase in replicon sizes. Apparently, also other factors of the pre-RC, namely CDC6 or MCM2-7, were not limiting in this process and the authors proposed that another, yet unidentified factor might be responsible for the observed changes.

The idea of a limiting factor that determines origin activation during S-phase, has been investigated many times and represents a suitable model to explain the successful propagation through S-phase from origins with variable firing efficiency (Goldar et al., 2008; Rhind, 2006). Many of the *trans*-acting factors are conserved between different species (Bogan et al., 2000) and a good candidate for a limiting factor is the CMG complex component CDC45, which has been found to be rate limiting for origin activation in different systems (Edwards et al., 2002; Mantiero et al., 2011; Tanaka et al., 2011).

Further molecular insight supporting this idea was provided by *Wong et al.* The authors determined the stoichiometry of factors important for licencing and activation of origins, similarly to what had earlier been published for the early developing *Xenopus* system (Blow et al., 2001; Edwards et al., 2002; Oehlmann et al., 2004; Rowles et al., 1996; Walter and Newport, 1997; Wohlschlegel et al., 2002). While CDC45 is present at levels that allow binding of 1-2 CDC45 molecules per replicon in *Xenopus* embryos, each of which is associated with 20-50 MCM hexamers, the combined results from hamster CHO, human and mouse fibroblast cells in their study indicated that the availability of CDC45 is limited to binding only every 600 - 1400 kb of DNA or ~2 CDC45 every six pre-replication complexes (Wong et al., 2011). It was further shown that CDC45, although present at lower levels in these cells, was stable and remained in the nucleus until the completion of S-phase. This suggests that differences in the availability of a limiting factors between embryonic and somatic stages controls the activation of replication origins. Indeed, many important cell cycle and replication factors are more abundant at the mRNA or protein level in embryonic stem cells of the mouse or are more stable over the course of S-phase in comparison to somatic cells, including cyclins, CDC45 mRNA and the initiation factor CDC6 (Fujii-Yamamoto et al., 2005).

The loading of MCM complexes in excess during the G1/S-phase, usually serves as a backup mechanism, that allows the activation of dormant origins, e.g. in the case of fork stalling events (Ge et al., 2007). However, usually >70% of those dormant origins remain silent (Moreno et al., 2016). In yeast, some of the key firing factors (Sld2, Sld3, Cdc45 and Dbf4) directly interact with the MCM helicase but are usually present at low levels compared to the frequency of licensed origins ((Douglas and Diffley, 2012) and references 13-16 therein). However, when available at increased protein levels, these and a few other factors induced the activation of both early and late activating origins (Mantiero et al., 2011; Tanaka et al., 2011).

Recently, it has been shown that the loading of MCM in human ES cells occurs at a very fast rate which was linked to their G1-phase, allowing stem cells to achieve similar levels of chromatin loaded MCM complexes than somatic cells with a longer G1-phase. Consistently, the rate of helicase loading slowed down with prolongation of the G1-phase upon differentiation (Matson et al., 2017). Embryonic stem cells of the mouse were even shown to load ~2-fold more MCMs onto DNA in comparison to differentiated neuronal progenitor cells (NPCs, (Ge et al., 2015)). The authors were able to count ~2500 MCM foci by 3D-SIM in a similar manner to the replication foci counting approach applied in this study. This number corresponds to roughly half the number of super-resolved replication sites that we

measured in this study, but the authors proposed that their MCM foci numbers might be an underestimation due to the resolution limit of the 3D-SIM technique. The actual dimensions of a MCM hexamer (25 x 16 nm, (Evrin et al., 2009)) therefore suggest that each focus contains multiple complexes and the total number of chromatin bound MCM might be even higher. Interestingly, in the work of *Fujii-Yamamoto* and co-workers, also the levels of phosphorylated MCM proteins were higher in mESC than in MEFs, reflecting increased activation of the replicative helicase during S-phase (Fujii-Yamamoto et al., 2005). This finding, together with the abundance of replication (initiation) factors in ES cells, could represent the molecular prerequisite necessary to explain an increased activation of replication origins in mammalian embryonic stem cells, in a manner similar to that of embryonic stages of frogs and flies.

The work of *Ge et al.* included also measurements of inter origin distances of origins within a cluster, that were on average even shorter (~50 kb) than what we had obtained from our samples (~90 kb). This could represent cell line specific differences or might be explained on the basis of the signals that were used to calculate the distances between neighboring origins. Although the authors did not compare this measurement with NPCs, such IOD is smaller than the values generally reported for somatic cells including our own measurements in mouse myoblasts and human cells (Berezney et al., 2000; Blumenthal et al., 1974; Chagin et al., 2016; Ge et al., 2007; Norio et al., 2005). This could indicate the existence of a further mechanism, by which the frequency of origins and the resulting replicon sizes are defined by a different organization of origins or origin clusters along chromosomes of embryonic over lineage committed cells, respectively.

Evidence for a molecular basis to explain such different organization is given by the work of *Guillou et al.*, who showed that activation of origins can be affected by differences in cohesin-mediated chromatin loop organization (Guillou et al., 2010). Cohesin was found to bind to replication origins in a manner independent of pre-RC assembly, which has also been observed in yeast and *Drosophila* (Glynn et al., 2004; Lengronne et al., 2004; MacAlpine et al., 2010). Knock-down of cohesin resulted in larger interfork distances and an overall decreased number of origin initiations, but did not affect replication fork speed. Interestingly, loss of cohesin did also not affect the overall number of replication foci (at conventional resolution) but only their intensity, which indicates that each focus contained less activated origins. Together with the observed changes in chromatin loop size it was proposed that down regulation of cohesin alters the spatial organization of replicons within

replication foci, with larger loops that contain fewer replication origins that results in longer replicons upon activation.

Although several studies already reported that the changes in the DNA replication program that occur during development are accompanied by variations in the spacing of origins as well as their initiation frequency, no systematic adjustments of replication fork speed were observed which could otherwise also influence the efficiency of DNA replication (Blumenthal et al., 1974; Liang et al., 1995; Walter and Newport, 1997; Wong et al., 2011). Nonetheless, a few studies suggest a tight connection between the rate of replication fork movement and origin activation. It was shown, for example, that differences in replication fork speed during replication directly affected the size of chromatin loops and the selection of origins for the subsequent S-phases (Zhong et al., 2013), while another study concluded that replication fork speed rather depends directly on the number of activated origins, i.e. forks would travel faster when only a low number of origins is fired and *vice versa* (Courbet et al., 2008). Thus, although the exact link between these two mechanisms remains unknown, it is possible that they are also interdependent on the availability of another rate limiting factors, i.e. nucleotide precursors for the synthesis of nascent DNA. In this, a higher activation of origins, which leads to more simultaneously active replication forks, could be limited by the availability of nucleotides from the cellular pool and, as a consequence, limit the rate of these increased number of active replication forks. The mouse myoblast cells used in the previous study also exhibited a significant increase in fork speed and the authors noted that this would have direct implications on the availability of nucleotides (Chagin et al., 2016). With regard to potential differences in origin activation between these cells and embryonic stem cells, as suggested in this work, this could even mean that these cells regulate their DNA replication dynamics on two different but likely related levels.

Regarding the discrepancy between calculated replicons and counted replication foci, I initially asked if this could be explained by a high proportion of unidirectional replication forks that might be present in embryonic stem cells. One well known example of a large DNA segment that gets replicated by a unidirectional replication fork, is the immunoglobulin heavy chain (IgH) region found in most mouse cells. In this locus, which is bounded by a ~250 kbp early-replicating segment on one and an at least 600 kbp region on the other side, a single replication fork travels from the early- to the late-replicating segment, thereby duplicating the whole 400 kbp IgH region positioned in the middle (Ermakova et al., 1999). Unidirectional replication forks are also one of two possible explanations for the existence of so called *timing transition regions* (TTRs), i.e. regions between constant early or late replication timing regions (replication domains, RDs), obtained from genome-wide

replication timing data (Rhind and Gilbert, 2013). In addition, unidirectional replication forks can arise as a results of fork stalling events, e.g. due to DNA damage sites that prevent fork passing, but there is no apparent biologically significant benefit in employing unidirectional replication forks in the context of normal genome duplication.

A recent study, that included mouse embryonic fibroblasts, showed that these cells exhibit around 5% unidirectional forks (Stanojcic et al., 2016). The signals obtained from our DNA combing datasets suggest an equally low level of one-sided replication forks in embryonic stem cells (data not shown). If the fraction would have been high enough (say for example 25%), it would have to be considered in the quotient of the calculation of simultaneously active replicons (Tab 4), as replicons are defined as the result of bidirectional replication forks starting from a single origin of replication. Consequently, if 25% of all measured replication forks were unidirectional, this would shift the ratio between replicons and nanoRFi to the same level as in mouse myoblast cells (i.e. 0.77), and would provide a reasonable explanation for the observed discrepancy. However, the fiber data it does not seem that a high percentage of unidirectional replication forks are the reason for the observed differences between replicons and replication nano foci in ES cells.

Although not further specified, *Chagin et al.* proposed that the lower ratio of replicons per super-resolved replication foci in the myoblast cells could be a result of single replication forks that are resolved by 3D-SIM. These signals would increase the total count of replication foci, thus influencing the calculations in a similar way.

In an attempt to measure individual replication forks to more directly address this possibility, I compared the fluorescence intensity of endogenous PCNA present at individual replication foci (Fig. 15A). PCNA is an important component of the replisome and we assumed that single replication forks could be discriminated from replicons (i.e. two replication forks) by differences in fluorescence intensity of PCNA per replication focus. This measurements, however, did not result in clear differences between the somatic and embryonic stem cells (Fig. 16 B). This may be attributed to the use of antibodies to detect the endogenous proteins which could bind in non-stoichiometric amounts to their epitope and thus influence the intensity based measurements negatively.

Although FRAP studies showed that the turnover of PCNA at the replication fork is very limited, with recovery times of more than 10 minutes (Sporbert et al., 2002), it is yet unclear how many PCNA molecules are acting simultaneously on the leading and lagging strand. Thus, varying quantities of the DNA clamp would make it impossible to measure differences

as little as the that between single and double replication forks. In addition, the two replisomes within a *replication bubble* could remain spatially associated throughout the duplication of a replicon and thus might not be measured independently from each other.

Another possibility that could affect the spatial positioning of replication forks or replicons is a different chromatin loop organization in mESCs at sites of ongoing DNA replication, e.g. by subdivision into multiple smaller loops. Like this we might also detect more individual forks that would result in even higher numbers of nano replication foci. The observed differences in heterochromatin compaction in embryonic stem cells (Fig. 11) are representative for a chromatin organization that is generally more open than that of differentiated cells (Lopes Novo and Rugg-Gunn, 2016; Meshorer and Misteli, 2006). This, in turn, could affect the organization of the DNA fiber at the lower level, i.e. into chromatin loops. A recently published study is in support of this, by providing insight in the organization of TADs into multiple sub-megabase sized domains (i.e. sub-TAD domains located within TADs) of several developmentally regulated genetic loci in mouse ES and neuronal progenitor cells (NPCs, (Phillips-Cremins et al., 2013)). Using the chromosome conformation capture carbon copy (5C) method, which provides high resolution of chromosomal contacts of genetic loci, the authors uncovered a range of cell type specific interactions within TADs of ESCs and NPCs. Although the study concentrated on the role of these interactions in the regulation of gene expression, similar differences could exist and play a role during DNA replication.

Measurement of the number of nanoRFi within a given nuclear volume, delimited by the same replication signals at lower resolution, however, resulted in the same number for mouse myoblast and mouse embryonic stem cells (Fig. 17 & Tab. S3). Although we cannot directly assess the conformation of the underlying chromatin fiber microscopically, this result indicates that the spatial organization or replicons does not differ substantially between somatic and ES cells. It remains possible, however, that the larger DNA content of myoblast cells results in more DNA within each of the analyzed volumes, and with that more nanoRFi. This would especially be the case if the myoblast cells, indeed, activate a higher number of replication origins than ES cells (Tab. 4).

Taken together, I conclude that with the approach published by *Chagin et al.*, it is possible to resolve the same functional units of DNA replication, i.e. replicons, in both lineage committed human, mouse myoblast and pluripotent mouse embryonic stem cells. The observed discrepancy between the experimentally determined number of active replicons and the microscopically resolved replication nano foci cannot be explained by a high

frequency of unidirectional or highly asymmetric DNA replication forks in mES cells. It remains possible that a large fraction of single replication forks, resolved as a results of the loose stem cell chromatin and the power of the 3D-SIM system, can account for this, although it could not be proven experimentally. Also, further experiments are required to specifically address any differences in the chromatin organization between pluripotent stem cells and differentiated cells that would allow a definite conclusion about an effect of such differences on the organization and regulation of DNA replication.

Finally, with regard to the differences in genome size, the data obtained from DNA fiber experiments (i.e. replication fork speed and inter origin distance) in combination with the cell cycle characteristics of the different cell types, may indicate interesting differences regarding the activation of replication origins during development in mammalian species.

DNA methylation could fine tune DNA replication but does not affect the global replication timing program in ES cells

The maintenance of DNA methylation at the C5 position of cytosines plays an important role in the regulation of gene expression, genomic imprinting and X-chromosome inactivation and is also tightly linked to DNA replication, during which existing DNA methylation patterns are maintained. The DNA methyltransferase DNMT1 has binding affinity for hemi-methylated DNA (Hermann et al., 2004) and contains a PCNA-binding motif that allows transient interaction with the DNA clamp during DNA replication (Chuang et al., 1997; Schermelleh et al., 2007), which was proposed to enhance the local concentration of DNMT1 to increase the establishment of newly methylated cytosines (Schermelleh et al., 2007). Consequently, loss of DNMT1 in mice causes embryonic lethality and somatic cells that lack DNA methylation are either strongly affected in cell proliferation or undergo apoptosis within a few cell divisions (Jackson-Grusby et al., 2001). ES cells, on the other hand, seem to be less sensitive to the loss of DNA methylation, as both human and mouse embryonic stem cells lacking all DNA methyltransferases (DNMT1/3A/3B) are viable. Nonetheless, some species specific differences still exist, based on the observation that hESCs deficient for only DNMT1 are rapidly undergoing apoptosis, while mESC can tolerate the loss of the maintenance DNMT (Liao et al., 2015; Tsumura et al., 2006).

Given the many important roles of DNA methylation during embryonic development, the regulation of gene expression and its implication in cancer development, I aimed to investigate the effect of loss of DNA methylation on the DNA replication dynamics in pluripotent stem cells. Interestingly, I did not observe global changes in the spatio-temporal order of the DNA replication program in DNMT triple-knockout or Np95$^{-/-}$ cells compared to their respective wild type (Fig. 20) All five replication patterns characterized in wild type cells, were also observed in both mutant cell lines. More specifically, also the duplication of pericentromeric heterochromatin took place at the first half of S-phase.

Pericentromeric heterochromatin is usually enriched for repressive epigenetic marks that include increased levels of DNA methylation, trimethylated histone H3 at lysine 9 (H3K9m3, (Almouzni and Probst, 2011; Lehnertz et al., 2003)) and hypoacetylated histones (Agalioti et al., 2002; Chen and Townes, 2000) and these marks are required for the establishment and maintenance of their condensed and silent state. With regard to DNA replication, an advanced replication timing of pericentromeric heterochromatin has been reported for mouse m5S cells treated with 5-azacytidine (Takebayashi et al., 2005). A similar result was obtained for mouse fibroblast cells that are deficient for DNMT1 (Casas-Delucchi et al.,

2012). In the latter study, however, the observed changes in replication timing could be attributed to changes in the histone acetylation levels that resulted in decompaction of heterochromatic clusters in these somatic cells. Chromatin decompaction upon loss of DNA methylation has indeed been observed multiple times (Lehnertz et al., 2003; Tamaru and Selker, 2001; Tariq et al., 2003) and can be explained by changes in the recruitment of histone-modifying complexes that usually interact with DNA methyltransferases (Jones et al., 1998; Nan et al., 1998; Wade et al., 1999; Zhang et al., 1999). DNMT1, for example, has been shown to recruit the histone methyltransferases G9a and SUV39H1 to methylated DNA that then catalyze the mono-, di- and trimethylation lysine 9 of histone H3, respectively (Klose and Bird, 2006). And members of the methylcytosine binding domain (MBD) protein family interact with histone deacetylases (HDACs, (Bird and Wolffe, 1999)), all of which normally cause chromatin compaction in the course of heterochromatization.

Tsumura and colleagues reported that, although global CpG methylation was lost in the DNMT1/3A/3B deficient mESCs, binding of the heterochromatin protein HP1 to chromocenters and H3K9 trimethylation were unaffected (Tsumura et al., 2006)). Interestingly, also histone acetylation levels and other histone methylation marks (H3K4m3) were similar in wild type and knockout cells.

Np95, on the other hand, is responsible for the recruitment of DNMT1 to hemi-methylated DNA during the replication of heterochromatin (Sharif et al., 2007) and was shown to bind to H3K9m3 via a tandem Tudor domain (Rottach et al., 2010) and to interact with the histone methyltransferase G9a and the histone deacetylase HDAC1 (Kim et al., 2009; Unoki et al., 2004). Consequently, at least in mouse fibroblast cells, loss of Np95 was shown to result in increased histone acetylation levels at pericentromeric heterochromatin (Papait et al., 2007).

Thus, in contrast to the reported effects of loss of DNA methylation on chromatin organization and compaction, the finding that the replication timing of pericentromeric heterochromatin is unaffected in mutant ES cell lines, suggests that chromocenters are already decompacted to a level at which changes in DNA methylation, histone acetylation and consequently chromatin decompaction do not influence their replication timing early in S-phase of embryonic stem cells.

With regard to the previously reported role of Np95 in heterochromatin replication in somatic cells (Papait et al., 2007), the data presented here further suggest a different role for Np95 in the replication of pericentromeric heterochromatin in pluripotent cells.

Interestingly, localization of Np95 to chromocenters during S-phase is similar in wild type ES and differentiated cells (F. Hastert, unpublished results), so how exactly Np95 might be involved in heterochromatin replication in stem cells, remains unclear.

Although loss of DNA methylation did not affect global DNA replication dynamics in mESC, I observed a clear decrease in the fraction of cells at the second stage of S-phase, at which pericentromeric DNA is duplicated, in asynchronous populations of DNMT deficient ES cells (Fig. 21 B). With a comparable S-phase duration (Tsumura et al., 2006) and a similar amount of replicating cells (Fig. 21 A), this result suggests that replication of the heterochromatic clusters is faster in cells lacking DNA methylation, which could be explained by either an increased activation of replication origins or changes in replication fork speed. In line with the latter assumption, treatment of the mouse m5S cells used by *Takebayashi et al.* with 5-azacytidine not only advanced the replication timing of pericentromeric DNA but also affected the rate of replication forks within these regions (Takebayashi et al., 2005) and the data obtained from DNA combing experiments (Fig. 22) further support the idea that loss of cytosine methylation increases DNA replication fork speed *in vivo.*

One possible explanation for the increased replication fork speed could be that the degree of cytosine methylation modulates the stability of the DNA double helix. This idea comes from *in vitro* experiments using *high resolution melting (HRM) curve* analysis, which showed that DNA containing differently modified bases, i.e. unmodified cytosine, 5-methylcytosine (5mC), 5-hydroxymethylcytosine (5hmC) or 6-methyladenine, exhibits significantly different melting temperatures. Hereby, DNA containing the lowest level of DNA methylation, or higher levels of 5hmC, exhibited the lowest melting temperature ((Lopez et al., 2012; Zhang et al., 2017a) and unpublished results). *In situ* evidence to proof this hypothesis comes from FISH experiments performed with wild type and DNMT1$^{-/-}$ mouse embryonic fibroblast cells. By measuring the hybridization efficiency of major satellite specific probes at increasing temperatures, detectable hybridization was observed already at lower temperatures in cells that lack DNMT1, and thus DNA methylation, but not wild type cells. In addition, accumulation of the single-strand binding protein RPA after treatment of the cells with the polymerase inhibitor aphidicolin (APH), occured at a much faster rate in the DNMT deficient cells (C. Rausch, unpublished results). We therefore argue that differences in base composition or cytosine base modifications are able to modulate the efficiency of DNA metabolic processes like transcription and DNA replication, that rely on local *melting* and continuous unwinding of the DNA double-helix.

OUTLOOK

As usual in science, the exciting details concerning the DNA replication dynamics in embryonic stem cells revealed in this work, lead to new and equally exciting questions that can and will be addressed in the future.

With regard to the duplication of pericentromeric heterochromatin early in S-phase in embryonic stem cells, we are currently analyzing and comparing a range of both important active and repressive post-translational modifications (PTMs) of histones between undifferentiated and differentiated ES cells. First of all, this allows us to compare potential differences in these modifications within the same cell lines, providing a better and direct insight to the changes that may occur during development. Furthermore, with this approach we aim to determine when during the differentiation process a *replication timing switch* occurs and if this can be attributed to specific epigenetic modifications or combinations thereof. As was shown in many previous studies, histone acetylation is one of the most promising candidates for the control of replication timing. Besides the more frequently used drug treatments to induce changes in PTMs, one could directly recruit either histone acetyltransferases or deacetylases to pericentromeric regions to increase or decrease acetylation levels more specifically, e.g. by employing molecular tools that are already available in our lab and that have been extensively used and characterized in one of our previous publications (Heinz et al., 2018). Beyond that, there are certainly many additional possibilities to assess whether the degree of chromatin compaction itself has an influence on the DNA replication timing of pericentromeric heterochromatin. Mass-spectometry analysis of chromocenters isolated from mES cells, for example, could help to figure out whether certain chromatin remodelling factors are specifically associated with or absent from pericentromeric heterochromatin of stem cells, respectively and whether changes in either direction, e.g. by overexpression or knock-down experiments, affect the compaction level of chromocenters. Another possibility could be the treatment of ES cells at different cell cycles phases with a hyperosmotic solution that was shown to induce global chromatin compaction and affected both the accessibility for and binding kinetics of histones and HP1 (Martin and Cardoso, 2010). Both experimental setups would allow to analyze both *in situ* and *in vivo* whether the replication timing of chromocenters is altered and maybe also whether it is controlled at the level of origin activation during S-phase or already at the licensing step that occurs during G1-phase. Especially in the latter scenario it would be very interesting to analyze the abundance of chromatin bound origin recognition complexes at pericentromeric heterochromatin. Given that ORC was shown to specifically interact with repressive heterochromatin marks (H3K9m3, H3K27m3 and H4K20m3 (Bartke et al., 2010;

Pak et al., 1997; Vermeulen et al., 2010)), a potential enrichment of ORC at chromocenters of mESC in combination with better accessibility for origin activator proteins would represent an exciting molecular basis for an earlier onset of replication. Changes in ORC levels, e.g. upon treatment with ORC specific inhibitors, could then help to answer the question whether advanced replication is controlled at the level of origin licensing. Another, more exotic experiment, could involve the isolation of chromocenters from murine somatic cell types (e.g. the mouse myoblasts presented in this work) and their transfer into mES cells. Labeling of the somatic cells with nucleotide analogues before their isolation, could allow to assess their replication timing upon injection into stem cells that are then labeled with another modified nucleotide. Simultaneous detection of both analogues could subsequently reveal if the somatic heterochromatin still gets replicated at a later time point, probably due their increased compaction and the enrichment of repressive histone and DNA modifications. This, however, would require detailed knowledge about the composition of pericentromeric heterochromatin of differentiated cells upon isolation, something that has been difficult to achieve in past experiments.

The differentiation of ES cells according to protocols used in many of the recent genome-wide studies offers further promising possibilities to address questions regarding differences in the rate of origin activation between pluripotent and somatic cells. To more specifically measure actual activation events in replicating cells, the most straightforward experiment to address this question would be to stain naïve and differentiated stem cells for endogenous phosphorylated MCM helicases (pMCMs, (Montagnoli et al., 2006)) and analyze their numbers by 3D-SIM. In regard of the phosphorylation events taking place at the replication fork during the onset of S-phase, phosphorylated helicases represent a good and more direct readout for activation of origins. As the number of active helicase molecules at each replication fork is known (two MCM helicases considering bidirectional replication), super-resolution foci measurements of pMCM would be a more accurate approach than for example the measurement of antibody-based fluorescence intensities of other replisome factors. Instead, one could also use the FACS-based approach presented by *Matson et al.*, who analyzed and compared the rate of MCM loading in hESCs versus neuronal precursors cells. The authors used a combination of EdU labeling, staining of chromatin bound MCM (i.e. after extraction of the unbound nuclear fraction) and DNA counterstaining to identify different cell cycle stages to quantify the amounts of MCM at each stage. Using a similar approach with antibodies specific for the phosphorylated, i.e. activated, MCM complexes over the course of S-phase could allow the measurement and comparison of activation between stem cells and somatic cells on a single-cell level. In additon to the assessment of differences in activation events during differentiation, it would

be equally interesting to also modify the level of potential factors that appear to be limiting for origin activation. First and foremost, this could be the CMG component CDC45. CDC45 overexpression, however, has been difficult if not impossible to achive in previous experiments and negatively affected the proliferative capacity of the transfected cells (personal communication and (Wong et al., 2011)). Given that we would rather expect higher amounts of CDC45 in embryonic stem cells (Fujii-Yamamoto et al., 2005) that could lead to increased origin activation, we would instead attempt to lower the protein levels CDC45 or related factors. If, indeed, activation of replication origins can be modulated this way, we could further evaluate related changes, e.g. in chromatin loop sizes and organization as well as replication fork speed, similar to what has been observed in previous studies.

One way to overcome the technical limitations that are imposed by the application of antibodies or overexpressed proteins, is by fluorescently tagging of endogenous proteins, e.g. with the help of the CRISPR/Cas system. This would help to avoid potential artifacts introduced as a result of protein overexpression as well as simpler problems, like the availability of specific antibodies for the intended target protein(s). By tagging endogenous PCNA, for example, we would probably be able to more directly address, whether single replication forks can be resolved by super-resolution microscopy techniques, instead of relying on signals from nascent DNA. In addition, this would allow us to address another long standig question, which is about the stoichiometry of PCNA at the replication fork. However, the addition of a tag is not always trivial and its usefulness, even if successful, is also dependent on the abundancy of the protein-of-interest within cells.

Collectively, based on the discussed data from previous publications and the results presented in this work, it becomes tempting to speculate that some of the developmental differences, with respect to DNA replication, seen in *Xenopus laevis* and *Drosophila melanogaster,* may also apply to higher mammalian species, including mice. Furthermore, these differences might be achieved and regulated by similar mechanisms, i.e. most likely changes in the activation of origins of DNA replication.

MATERIALS & METHODS

Cell Culture

All embryonic stem cell lines were maintained at 37 °C and 5 % CO_2 in DMEM high glucose (Cat.No.: D6429, Sigma-Aldrich Chemie GmbH, Steinheim, Germany) supplemented with 16 % FCS, 1x non-essential amino acids (Cat.No.: M7145, Sigma, Germany), 1x Pen/Strep (Cat.No.: P4333, Sigma-Aldrich Chemie GmbH, Steinheim, Germany), 1x L-glutamine (Cat.No.: G7513, Sigma-Aldrich Chemie GmbH, Steinheim, Germany), 0.1 mM beta-mercaptoethanol (Cat.No.: 4227, Carl Roth, Karlsruhe, Germany), PD 0325901 (0.1 µM, Cat.No.: Axon 1408, Axon Medchem BV, Groningen, The Netherlands), CHiR 99021 (0.3 µM, Cat.No.: Axon 1386, Axon Medchem BV, Groningen, The Netherlands), LIF (1,000 U/ml; Cat.No.: ALX-201-242, Enzo Life Sciences GmbH, Lörrach, Germany) on gelatin-coated culture dishes (0.2% gelatin; Cat.No.: G2500, Sigma-Aldrich Chemie GmbH, Steinheim, Germany). Culture medium was changed every day and cells were split every two days.

HeLa (ATCC CCL-2), male mouse embryonic fibroblast (MEF W8, (Lander et al., 2001)) and C2C12 mouse myoblasts (ATCC CRL-1772) cells were cultured in cell culture medium (DMEM, Cat.No.: D6429, Sigma-Aldrich Chemie GmbH, Steinheim, Germany) containing the appropriate concentrations of FCS (10% for HeLa and MEF W8 cells, 20% for C2C12, respectively), supplemented with 50 µg/ml gentamycin (Cat.No.: G1397, Sigma-Aldrich Chemie GmbH, Steinheim, Germany) and 20 mM L-glutamine (Cat.No.: G7513, Sigma-Aldrich Chemie GmbH, Steinheim, Germany) at 37°C, 5% CO_2.

For maintenance cells were washed with sterile 1x PBS/EDTA followed by trypsinization using trypsin/EDTA solution (Cat.No.: T4049, Sigma-Aldrich Chemie GmbH, Steinheim, Germany). Trypsin reaction was stopped by adding fresh cell culture medium and cells were seeded as required.

A list of all mouse embryonic stem cell lines used in this study is provided in Table 5.

Table 5 - mES celll lines used in this study

Cell Line	Cell Type	Sex	Reference
J1	mESC	male	(Li et al., 1992)
J1 Dnmt1$^{-/-}$/Dnmt3a$^{-/-}$ /Dnmt3b$^{-/-}$ triple KO	mESC	male	(Tsumura et al., 2006)
E14	mESC	male	(Doetschman et al., 1987)
E14 Np95 $^{-/-}$	mESC	male	(Muto et al., 2002)

Pulse-chase labeling

Cells were incubated with cell culture medium containing 10 µM EdU final concentration (Cat.No.: C10085, Invitrogen, Karlsruhe, Germany) for the time indicated and chased for varying lengths at 30 min increments with fresh medium supplemented with thymidine (50 µM final concentration), followed by fixation. EdU detection was performed first, followed by antibody detection as described above. EdU detection by ClickIT chemistry was performed according to the manufacturer's instructions (Invitrogen, Carlsbad, CA, USA).

Immunofluorescence

For immunostaining experiments cells were grown on glass coverslips. For embryonic stem cells, coverslips were coated with 0.2 % gelatin (Cat.No.: G2500, Sigma-Aldrich Chemie GmbH, Steinheim, Germany). Cells were fixed with 3.7 % formaldehyde (Cat.No.: F8775, Sigma-Aldrich Chemie GmbH, Steinheim, Germany) buffered with 1x PBS for 10-15 minutes. After washing with PBS-T (1x PBS/0.02 % Tween 20), cells were permeabilized using 1x PBS/0.5 % Triton-X100 for 12-15 minutes, washed again with PBS-T and blocked with blocking buffer (1x PBS/0.02% Tween/2% BSA; (Kraus et al., 2017)) for 30 minutes.

For detection of endogenous PCNA cells were further incubated for 5 minutes in ice-cold methanol on ice following formaldehyde fixation to ensure antigen accessibility. Methanol was exchanged step-wise and cells were finally washed 2-3 times with PBS-T before the blocking step.

Primary and secondary antibodies were diluted in blocking buffer and incubation was performed for 45 - 60 minutes at 37 °C in a humidified staining chamber covered with aluminum foil.

In case of BrdU detection, antibodies were diluted in buffer consisting of a 1:1 mixture of blocking and 2x DNase I reaction buffer (60 mM Tris/HCl pH 8.1, 0.66 mM $MgCl_2$, 1 mM beta-mercapthoethanol). DNase I (2 U/µl, D5025, Sigma-Aldrich Chemie GmbH, Steinheim, Germany) was included at a final concentration of 0.6 U/reaction. Samples were incubated for 1 hour at 37 °C and DNase I digestion was inhibited by washing the coverslips with PBST-T/EDTA (1 mM EDTA). EdU detection was performed according to manufacturer's instructions.

Antibodies and reactive azides used for fluorescent detection are listed in Table 6

Table 6 - List of antibodies and fluorescent azides used in this study

Antigen	Host	Clon-ality	Company	Cat. No.	Lot	Dilution
Primary Antibodies						
anti PCNA	ms[1]	mAb	DABCO	M0879	00083603	1/200
anti BrdU	rat	mAb	Biorad	OBT0030CX	0714	1/100
anti BrdU	ms	mAb	BD	347580	2205877 & 29-07-13	1/200
anti single-stranded DNA (IgG2a)	ms	mAb	Millipore	MAB3034	2684913	1/200
anti Cytosine(5mC)	ms	mAb	Eurogentec	MMS-900P-B	10110	1/100
anti Oct3/4	ms	mAb	BD Bio-sciences	611203	5079562	1/1000
Secondary Antibodies						
anti mouse Chromeo 546	gt[2]	pAb	Active Motif	15033	26709005	1/600
anti-mouse IgG2a AlexaFluor 647	gt	pAb	Fisher Scientific GmbH	A-21241	1755551	1/300

anti rat IgG (H+L) AlexaFluor 488	dk[3]	pAb	Jackson Immuno Research	712-545-153	113974	1/500
anti mouse IgG (H+L) AlexaFluor 594	dk	pAb	Jackson Immuno Research	715-585-151	127802	1/500
Add. Reagents						
Streptavidin AlexaFluor 488 conjugate	-	-	Invitrogen	S11223	00083603	1/200
Alexa Fluor 488 Azide	-	-	Invitrogen	C10337 Comp. B		
Alexa Fluor 594 Azide	-	-	Invitrogen	C10330 Comp. B		
6-FAM azide	-	-	Carl Roth	7806	201112008	1/500

[1] = mouse; [2] = goat; [3] = donkey

Live Cell Microscopy

For live cell experiments J1 mESCs were co-transfected with plasmids encoding for mCherry-tagged PCNA (Sporbert et al., 2005) and eGFP-tagged major satellite-specific zinc finger binding protein (MaSat-GFP, (Lindhout et al., 2007)) using AMAXA Nucleofection (Amaxa Nucleofector II, Lonza Ltd., Basel, Switzerland) or poly-ethylenimine (PEI, 1 mg/mL in ddH_2O, pH 7; Sigma Aldrich, St. Louis, MO, USA). After transfection, cells were plated on p35 dishes containing a glass bottom that allowed for optical imaging. Time-lapse imaging was performed with a UltraVIEW VoX spinning disc system (Perkin Elmer, Massachusetts, USA) mounted on a Nikon Ti microscope using an oil immersion objective lens Plan-Apochromat x60/1.45 NA in a closed live-cell microscopy chamber (ACU control, Olympus) at 37 °C, 5% CO_2 and 60% humidity. Z-stacks were acquired with a Hamamatsu C9100-50 EMCCD camera to analyze whole nuclear volumes.

Molecular combing

Molecular combing experiments have been performed as described before (Bialic et al., 2015). Briefly, cells were pulse labeled for 15 minutes with 25 µM 5-iodo-2'-deoxyuridine (IdU) first. After washing twice with pre-warmed PBS or cell culture medium incubation with 5-chloro-2'-deoxyurdine (CldU, 200 µM) pulse of the same length followed. Cells were subsequently embedded in low-melting point agarose (NuSieveGTG agarose, Biozym Scientific GmbH, Hessisch Oldendorf, Germany), genomic DNA was isolated by proteinase

K (final concentration: 0.2 mg/ml diluted 100x from 20 mg/ml stock) digestion and single DNA molecules were stretched on vinylsilane□coated glass coverslips (Genomic Vision, Bagneux, France). Incorporated nucleotides were visualized by immunofluorescence using the following antibodies: mouse anti-BrdU (Becton Dickinson, cat. No. 347580 , Heidelberg, Germany, diluted 1:200) to detect IdU, rat anti-BrdU (AbD Serotec, cat. No. OBT0030CX, Puchheim, Germany, diluted 1:200) for the detection of CldU, mouse IgG2a anti-single stranded DNA (Millipore, cat. No. MAB3034, Darmstadt, Germany, diluted 1:200), anti-mouse IgG Chromeo 546 (Active Motif, cat. No. 15033, La Hulpe, Belgium, diluted 1:300), anti-rat IgG AlexaFluor 488 (Jackson Immunoresearch, cat. No. 712-545-153, Suffolk, UK, diluted 1:300), anti-mouse IgG2a AlexaFluor 647 (Fisher Scientific GmbH, cat. No. A-21241, Schwerte, Germany, diluted 1:300).

Sufficient replication signals were measured using ImageJ (http://imagej.nih.gov/ij/) and replication fork speed was calculated by dividing the track length (in kilo base pairs) by the time of nucleotide application.

Replication foci analysis by 3D-SIM

For the analysis of replication foci (RFi) numbers in fixed samples, cells were grown on high-precision cover glasses (Carl Roth, Cat.No.: LH22.1, 18 x 18 mm (170 ± 5 µm); #1.5H) and subsequently treated and stained as described above. In case of ESCs, cells were re-plated several hours before labeling and fixation to achieve single cells suitable for imaging. This procedure did not result in measurable effects on the spatio-temporal order of replication patterns.

Samples were imaged with a DeltaVision OMX V3 system (GE Healthcare) equipped with a 100x/1.40 NA PlanApo oil immersion objective (Olympus). Available laser lines are 405, 488 and 593 nm diode lasers.

A protocol for the quantification of replication foci numbers within individual cell nuclei as developed in our lab has been published in (Chagin et al., 2015). In brief, raw 3D-SIM images were first converted to 16-bit images. Further image pre-processing was performed in ImageJ and involved the segmentation of individual cell nuclei using maximum intensity projections of the DAPI signal. In a next step replication signals were segmented by auto-thresholding using the *Triangle* method. The resulting binary images were used to mask the original replication foci signals of interest and to discriminate them from background (set to "0"). These images were imported to the image analysis software Volocity 6.3 (Perkin

Elmer) and replication foci were quantified for individual nuclei. 3D-SIM replication foci were detected by intensity excluding only black pixels (i.e. background with intensity "0"). Additionally, touching foci were separated (Object size guide = 0 μm^3) and signals smaller than 0.002 μm^3 were excluded from the final counting as they represented unspecific background signal (evaluated manually).

For the analysis of replication forks per individual replication focus (Fig. 16) C2C12, HeLa and J1 mESCs were labeled for 20 min with 20 µM BrdU and immediately fixed. Experiments were performed in triplicates and immunodetection of BrdU and PCNA signals as well as image analysis was performed as described above. Imaging settings for 3D-structured illumination microscopy were the same for all samples. Measurement of PCNA intensities within segmented BrdU nano-foci was restricted to signals with sufficient overlap. Previous measurements in which we aimed to define signal colocalization for super-resolution images (Scholl *et al.*, in preparation) showed that a maximal centroid distance of two signals below 0.1 µm is sufficient to consider colocalization. Based on this assumption, an additional filtering step was included in the analysis pipeline that excluded those foci were the centroids of the individual signals were more than 0.05 µm apart from each other.

Pseudo wide-field (pseudoWF or pWF) replication signals were obtained from the same datasets (C2C12, HeLa and J1 mESCs with BrdU). Generation of the pseudo wide-field data was described in (Natale et al., 2017). For correlation analysis the pseudoWF images had to be processed initially using ImageJ to match the image dimension of the 3D-SIM data. First, the images were scaled using a bicubic interpolation, doubling the number of pixel in x and y. This resulted in similar voxel sizes between the pseudoWF and 3D-SIM datasets (40 x 40 x 125 nm). Next, pseudoWF images were corrected for slight pixel shift by translating the image stack -2 pixels in x and y direction (ImageJ > *Image* > *Transform* > *Translate*). For segmentation of the pseudoWF replication signals, the histogram was normalized (ImageJ > *Process* > *Enhance Contrast* > Saturated pixels = 0%, check normalize histogram*)* and a background substraction was performed using a rolling ball algorithm with radius = 10 (ImageJ > *Process* > *Substract Background* > Radius = 10*).* Segmentation was performed by auto-thresholding using the Otsu algorithm (ImageJ > *Image* > *Adjust* > *Auto-Threshold* > *Otsu; ignore black, ignore white, use stack histogram)*. 3D-SIM replication signals were processed as described above. Finally, segmented and masked pseudoWF and 3D-SIM image stacks were merged and used for foci counting in Volocity. Detection of pWF RFi was based on intensity as for 3D-SIM images and separation of touching objects was based on object size (Object size guide = 0.02 μm^3). Signals smaller than 0.02 μm^3 were excluded from the final counting. Overlapping signals

used for nanoRFi counting within pWF RFi were filtered by an additional compartmentalization step (*Compartmentalize* > *Divide items in: nanoRFi ; between items in: RFi; where sub-populations are: overlapping*).

Analysis of heterochromatin compaction

3D-SIM images of DAPI stained nuclei of J1 mESC and C2C12 myoblast cells were imported into Volocity 6.3 (Perkin Elmer). Pericentromeric heterochromatin (chromocenters) were segmented based on fluorescence intensity and the quality of the results was evaluated manually.

To compare the degree of compaction, the surface area as well as the shape factor were derived from the measurements for each chromocenter within each individual cell nucleus. The shape factor is given by the formula: Shape Factor = $\frac{\pi^{\frac{1}{3}}(6V_0)^{\frac{2}{3}}}{A}$ for a 3-dimensional sphere, which represents the ratio of the surface area of a sphere (same volume as the object of interest) to the surface area of the object of interest. A perfect sphere would results in a shape factor of 1, while a smaller value represents increasing deviation from the spherical structure. Data analysis and plotting was done in *RStudio* (v1.0.143).

Fluorescence in-situ hybridization

The method for fluorescence in-situ hybridization was derived from a protocol published by (Cremer et al., 2008).

Parts of the description of this method were taken from *Weber, P., Rausch, C, Scholl, A., Cardoso, M.C.: Repli-FISH: Application of 3D-(Immuno)-FISH for the study of DNA replication timing of genetic repeat elements* currently under revision and are marked accordingly.

The following sections were modified from *Weber et al.* and were written by C.R. and P.W.

Bio-16-dUTP nucleotide generation *(written by C.R.)*

Bio-16-dUTP labeling reactions contained 4.5 mM aminoallyl-dUTPs (5-(-3-aminoallyl)-2'-deoxyuridine 5'-triphosphate, Cat.No: A 0410, Sigma-Aldrich Chemie GmbH, Steinheim, Germany), 44.5 mM carbonate-bicarbonate buffer with 0.5 M NaCl (0.2 M NaHCO3 stock with 0.5 M NaCl titrated to pH 8.3 with 0.2 M Na2CO with 0.5 M NaCl) and 8.9 mM biotin-XX (Cat.No: B-1606, Molecular Probes) in a total volume of 45 µl. Reactions were incubated 3-4 hours at room-temperature and stopped by the addition of 20 mM glycine (pH 8.0). 20 mM Tris-HCl (pH 7.75) was added to stabilize the nucleotides.

Table 7 - *Primers used for FISH probe generation*

Primer Name	***Sequence [5' - 3']***	***Reference***
MaSat-F	*AAAATGAGAAACATCCACTTG*	*(Frauer et al., 2011)*
MaSat-R	*CCATGATTTTCAGTTTTCTT*	*(Frauer et al., 2011)*
MiSat-F	*CATGGAAAATGATAAAAACC*	*(Lehnertz et al., 2003)*
MiSat-R	*CATCTAATATGTTCTACAGTGTG*	*(Lehnertz et al., 2003)*
Telo-F	*TTAGGGTTAGGGTTAGGGTTAGGGTTAGGG*	*(Anton et al., 2014)*
Telo-R	*CCCTAACCCTAACCCTAACCCTAACCCTAA*	*(Anton et al., 2014)*

Major and minor satellite repeat specific probe *(written by P.W.)*

Both major and minor satellite repeat specific probes were generated via PCR amplification using repeat specific primers and mouse genomic DNA from C2C12 mouse myoblasts as template. Since this amplification results in probes of varying length (detectable as a smear on an agarose gel), samples need to be digested using DNase I (1:250 dilution in ddH2O (stock: 2U/µl), 2 µl/reaction) for a maximum of 30 min at room temperature. Addition of 1 µl 0.5 M EDTA stopped the reaction and probe quality and size were checked via agarose gel electrophoresis (1% agarose gel, 5 µl of reaction).

Table 8 - *PCR reaction setup for minor and major satellite repeat probes*

PCR component	*PCR (final concentration*
PCR buffer 0.08 mM	*1x Taq buffer**
dATP/dGTP/dCTP	*0.2 mM each*
bio-16-dUTPs	*0.08 mM*
primer F/R	*0.2 µM*
polymerase	*1-2 U Taq*
DNA template	*1 µl mouse gDNA*
Final volume	*to 100 µl with ddH2O*

**10x Taq buffer: 100 mM Tris/HCl pH 8.3, 500 mM KCl, 15 mM MgCl2.*

Table 9 - *PCR cycle conditions for minor and major satellite repeat probe amplification*

Cycle Step	*Temperature*	*Time*	*Cycle*
Initial Denaturation	*95 °C*	*5 min*	*1*
1. Denaturation	*95 °C*	*1 min*	*40*
2. Annealing	*56 °C*	*1 min*	
3. Extension	*72 °C*	*2 min*	
Final Extension	*72 °C*	*5 min*	*1*

Telomere repeat specific probe *(written by P.W.)*

Probes specific for telomeric repeats were generated by a two-step PCR using the self-annealing primers described in (Anton et al., 2014). The first PCR reaction served to create elongated probe fragments with regular dNTPs, while the second reaction included the biotinylated nucleotide analogues for efficient probe labeling. For the initial amplification a high-fidelity polymerase was used (e.g. Q5 polymerase (NEB, Cat#: M0491)). For labeling with bio-16-dUTPs in the subsequent PCR, however, Taq polymerase is required since the Q5 polymerase does not incorporate dUTPs. Cycle conditions remained the same in both reactions.

Table 10 - *PCR reaction setup for telomere specific probes*

PCR component	***PCR (final concentration)***	***Labeling (final concentration)***
PCR buffer	*1x Q5 buffer*	*1x Taq buffer**
$MgCl_2$	*0.5 mM*	*-*
dATP/dGTP/dCTP	*0.2 mM each***	*0.05 mM each*
dTTP	*0.2 mM***	*0.01 mM*
bio-16-dUTPs	*-*	*0.1 mM*
primer F/R	*0.2 µM each*	*0.2 µM each*
polymerase	*1-2 U Q5 DNA polymerase*	*1-2 U Taq*
DNA template	*-*	*1 µl of PCR product*
Final volume	*to 50 µl with dH_2O*	*to 50 µl with dH_2O*

**10x Taq buffer: 100 mM Tris/HCl pH 8.3, 500 mM KCl, 15 mM MgCl2.*

*** here a dNTP mix containing all four nucleotides was used.*

Table 11 - *PCR cycle conditions for telomere specific probe amplification*

Cycle Step	*Temperature*	*Time*	*Cycle*
Initial Denaturation	*95 °C*	*5 min*	*1*
1. Denaturation	*98 °C*	*10 s*	*40*
2. Annealing	*50 °C*	*15 s*	
3. Extension	*72 °C*	*20 s*	
Final Extension	*72 °C*	*2 min*	*1*

Probe precipitation and fluorescence in situ hybridization (FISH)

(This section was written in collaboration by C.R. and P.W.)

Probe precipitation

Probes were ethanol-precipitated with 0.15 M sodium-acetate and 50 µg/mL fish sperm DNA, washed with 70% ethanol, air-dried and resuspended in hybridization solution with formamide (70% formamide, 2x SSC (30 mM Na-citrate, 300 mM NaCl), 10% dextran sulfate, pH 7) for MaSat, MiSat and telomere probes and without formamide (10 mM Tris-HCl, 3 mM MgCl2, 50 mM KCl, 10 µg/mL gelatin, 2x SSC, (Celeda et al., 1994; Celeda et al., 1992)) for Alu and LINE-1 probes. Probes were denatured for 5 min at 80 °C and immediately put on ice.

Cell preparation

For subsequent hybridization and detection of replication signals cells were washed twice with 1x PBS and immediately fixed with 3.7% formaldehyde/PBS for 10-15 min. After washing 2-3x in PBS-T (0.02% Tween) cells were permeabilized in 0.5% Triton-X100/PBS for 10-15 min and washed again with PBS-T. At this point, cells may be stored for a few days in 1x PBS at 4 °C or used directly for hybridization. For MaSat, MiSat and telomere probe hybridization, cells were incubated for 15 min in 0.1 M HCl, washed 3x with 1x PBS and equilibrated first for 2 min in 2x SSC and finally for 20 min in 50% formamide/2x SSC. For Alu and LINE-1 FISH, cells were equilibrated in hybridization solution without formamide.

Hybridization

For hybridization a circle/square the size of the cover glass carrying the cells was drawn onto a regular microscope glass slide (Menzel-Gläser Superfrost, Thermo Scientific, Cat.No.: ABAA000080##32E) using an ImmunoPen (Calbiochem, Cat.No.: 402176-1EA). Next, the hybridization solution containing the respective probe was applied to the center of this mark and cells were applied face down onto the hybridization solution. The cover glass was additionally sealed with a silicon ring/square (selfmade) and a second microscope slide was applied on top. The whole apparatus was then transferred to a hybridization chamber and tightly sealed.

The chamber was subsequently incubated for 5 min at 80 °C in a water bath to denature both genomic DNA and probe, followed by incubation in a second water bath for 24 hours at 37 °C (MaSat, MiSat and telomeres) or 42 °C overnight (Alu and LINE-1).

Washing and blocking

After hybridization cells were carefully removed from the microscope slide and washed for 3x 5 min with 50% formamide/2x SSC pH 7 at 45 °C followed by a 2x 5 min wash in 2x SSC at 45 °C. For Alu and LINE-1 probes, the washing steps were performed with 2x SSC and 0.1x SSC without formamide. To block unspecific antibody binding sites, cells were incubated for 30 min with 1% BSA/4x SSC at room temperature.

***NOTE:** If PCNA detection is desired, immunodetection needs to be performed before continuing with the probe detection.*

Probe detection

Probes were detected by incubation with fluorescently-tagged streptavidin diluted in 1% BSA/4x SSC for 20 min at room temperature and washed 3x with 0.05% Tween/4x SSC for 5 min. A post-fixation step with 1-4% formaldehyde/PBS for 10 min further stabilized resulting antibody complexes. DNA counterstaining was performed by incubating cells for 10 min with DAPI (10 µg/ml, 4',6- diamidin-2-phenylindol) at room temperature and slides were mounted using Mowiol.

Replication staining ***(written by P.W.)***

Staining of the replisome component PCNA

Immunofluorescent detection of PCNA usually requires additional fixation with methanol to guarantee epitope accessibility. However, we noticed that this step can be omitted when the hybridization procedure has been performed before. Antibody detection of PCNA needs to be performed after hybridization and washing but before detection of the biotinylated probe to achieve good signals. After blocking as described above, cells are incubated for 1-2 hours at room temperature with an antibody specific for PCNA (mouse anti-PCNA, mAb, DABCO, Cat#: M0879, 1/200) diluted in blocking buffer (1x PBS/0.02% Tween20/2% BSA). After washing 2-3x in PBS-T (0.02% Tween20), detection with secondary antibodies conjugated with fluorophores suitable for subsequent microscopical detection was performed, before continuing with the probe detection.

DOP-PCR for the generation of Y chromosome specific probe template stocks

Y chromosome-specific template DNA was kindly provided by Prof. Dr. Diane Krause (Yale University School of Medicine). A template stock was generated via DOP-PCR (*degenerated oligonucleotide-primed-PCR*) from the DNA starting material using the 6AI primer:
5’ – CCGACTCGAGNNNNNNTACACC – 3’.

PCR reactions contained 2 µl mouse Y-chromosome template DNA, 1x PCR buffer (10 x stock: 100 mM Tris/HCl pH 8.3, 500 mM KCl, 15 mM MgCl2), 6AI primer (final conc.: 2 µM), dNTPs (final conc.: 0.25 mM) and 2.5 U Taq polymerase (selfmade) in a total volume of 50 µl.

Table 12 - *DOP-PCR reaction condition for Y-chromosome probe amplification*

Cycles	Denaturation	Annealing	Extension
1 - 2	45 sec at 94 °C	45 sec at 15 °C	12 min at 37 °C
5	40 sec at 94 °C	45 sec at 37 °C	4 min at 66 °C
24	40 sec at 94 °C	45 sec at 54 °C	4 min at 66 °C

Label DOP-PCR for the generation of biotinylated Y chromosome probes

Y chromosome template DNA was subsequently used in a second DOP-PCR reaction including biotinylated nucleotides.

Template DNA was mixed with a nucleotide mixture containing unlabeled (stock concentration: 0.5 mM each dATP, dCTP and dGTP with 0.1 mM dTTP) and biotinylated nucleotides (stock concentration: 1 mM biotin-16-dUTPs). 6AI primer, Taq polymerase and PCR buffer were used as in the first PCR. Total reaction volume was 50 µl.

Table 13 - *Label DOP-PCR reaction condition for Y-chromosome probe amplification*

Cycles	Denaturation	Annealing	Extension
1	5 min at 94 °C	-	-
35	30 sec at 94 °C	30 sec at 54 °C	90 sec at 72 °C
1	-	-	5 min at 72 °C

Karyotype analysis and genome size estimation

Metaphase spreads were prepared from J1 ES cells for karyotype analysis. Cells were plated on gelatine coated cell culture plates. To arrest cells in mitosis, colcemid was added to the cell culture medium at a final concentration of 0.02 µg/ml (stock concentration 10 µg/ml. Cat.No.: 10 295 892 001, Roche Diagnostics GmbH) and cells were incubated for for 1 - 1.5 hours at 37 °C. After collecting the supernatant and harvesting the plate by trypsinization cells were pelleted by centrifugation for 5 minutes at 300 x g. Supernatant was removed carefully and cells were resuspended in 10 ml pre-warmed hypotonic solution (0.075 M KCl) and incubated for 6 minutes at 37 °C. After centrifugation for 5 minutes at 300 x g cells were fixed by dropwise addition of fixative solution (3:1 methanol:acetic acid) and incubation for 45 minutes on ice. Etched microscope slides were prepared by submerging the slides for 15-20 minutes in etching solution (0.1 N HCl in 95% ethanol) followed by cleaning steps in 95% EtOH and ddH_2O (3 times in each solution). Finally, spreads were generated by dropping fixed cells onto etched slides and air-drying. Individual metaphase spreads were imaged by phase-contrast microscopy and analyzed manually.

Cell cycle profiling

Cell cycle profiles of J1 mES, C2C12 mouse myoblasts and human HeLa cervix carcinoma cells were compared by FACS profile analysis after propidium iodide (PI) staining as described in (Darzynkiewicz et al., 2017). Briefly, cells were collected by centrifugation for 5 minutes at 300 x g and supernatant was removed carefully. Cells were then resuspended in 1x PBS to achieve a single-cell suspension. Cells were fixed by drop-wise addition to 70% ethanol while shaking, followed by incubation for >2 hours. DNA staining was performed by overnight incubation of cells in PI staining solution (final concentrations: 1x PBS, 0.1% (v/v) Triton X100, 0.2 mg/ml RNase A and 20 µg/ml propidium iodide). Unstained control cells were incubated in a equal solution without PI. Experiments were performed in quadruplicates for all cell lines. Cell cycle profiles for all cell lines were acquired under the same conditions using a BioRad S3 Cell Sorter (488 nm 100 mW DPSS laser line).

Doubling time and S-phase duration analysis

Population doubling times were obtained from growth curve analyzes after counting cells with a Neubauer haemocytometer and a *Scepter 2.0 Cell Counter* (Merck Millipore), respectively for up to two weeks over the course of this study.

The approximate S-phase duration of J1 mESCs was derived from the population doubling time. First, the fraction of replicating cells within an asynchronous population was quantified microscopically in replicate experiments. The percentage of this sub-population was then used to calculate the length of S-phase from the overall cell cycle duration.

Image and statistical analysis and data visualization

Image analysis was done using ImageJ (http://rsb.info.nih.gov/ij/, v1.51s and earlier) using custom written scripts available upon request.

Statistical analysis and data visualization was performed with *RStudio* (v1.0.143) and *GraphPad Prism* (v5.0a for Mac, GraphPad Software, La Jolla California USA, www.graphpad.com), respectively. Final images were composed in Adobe Illustrator.

SUPPLEMENTARY MATERIAL

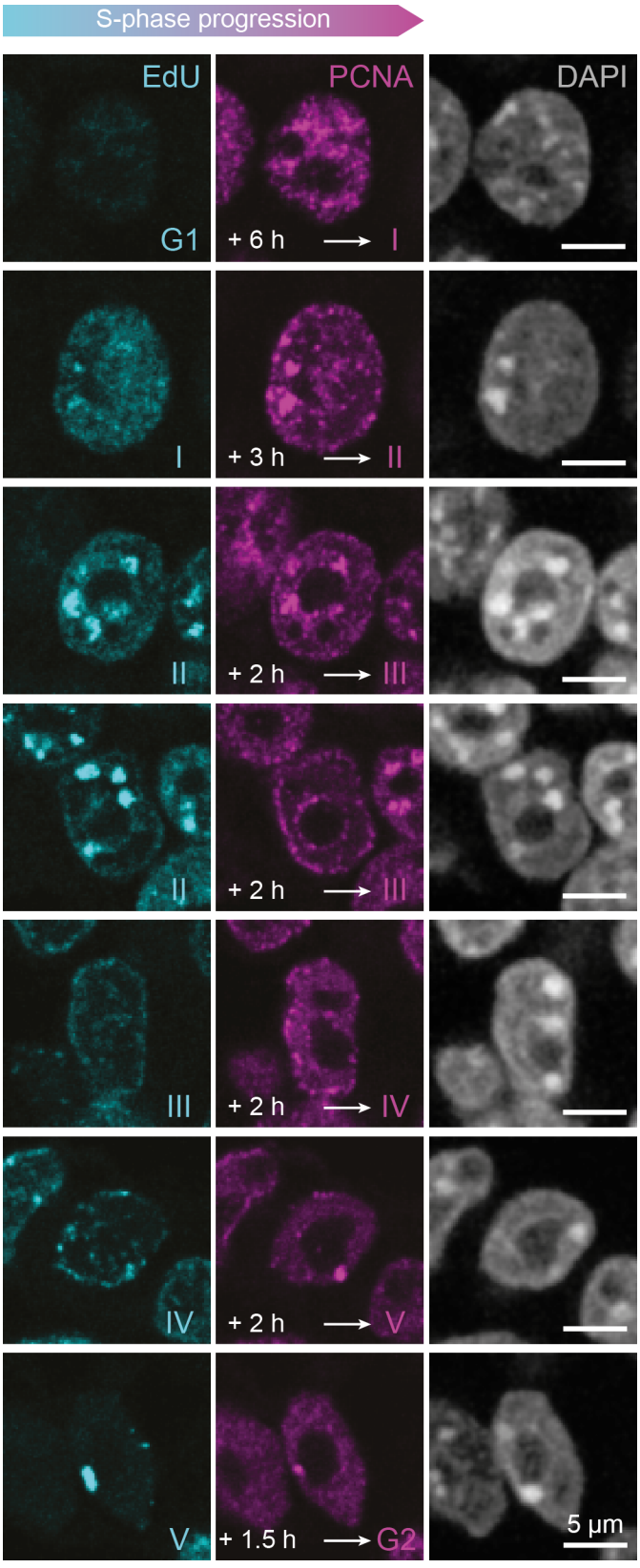


Figure S1 – The spatio-temporal order of DNA replication in ES cells is confirmed in a second cell line.
Male mouse E14 embryonic stem cells were labeled with EdU and stained similar to the J1 ES cell line shown in Fig. 5. EdU (cyan) marks the sites of DNA replication at the time of nucleotide application. PCNA (magenta) marks the subsequent pattern after the cells progressed for a certain amount of time (indicated in the images) through S-phase. Transition from one pattern to the next (indicated by G1, I-V and G2, respectively) can be observed by comparing the images from left to right in each row. All patterns characterized in J1 ESCs were observed in the E14 cell line as well.

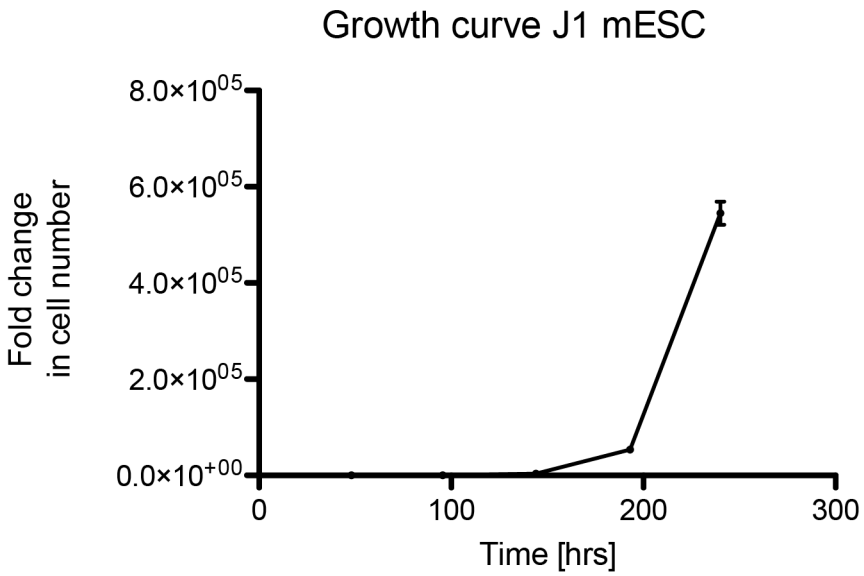


Figure S 2 - Growth curve of J1 mES cells.
The growth curve shows the fold change in cell number of asynchronously growing J1 ES cells over a time period of 240 hours or 10 days. From this duplicate experiment we calculated a doubling time of 12 hours.

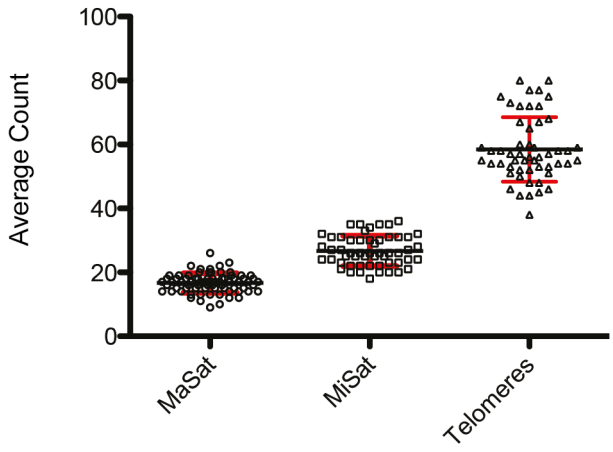


Figure S 3 - Manual quantification of tandem repeat signals in mES cells.
FISH signals of telomeres, minor and major satellite tandem repeats were assessed by-eye. Clustering of the individual elements is indicated by a lower than the expected average number of signals considering the diploid karyotype of J1 ES cells.

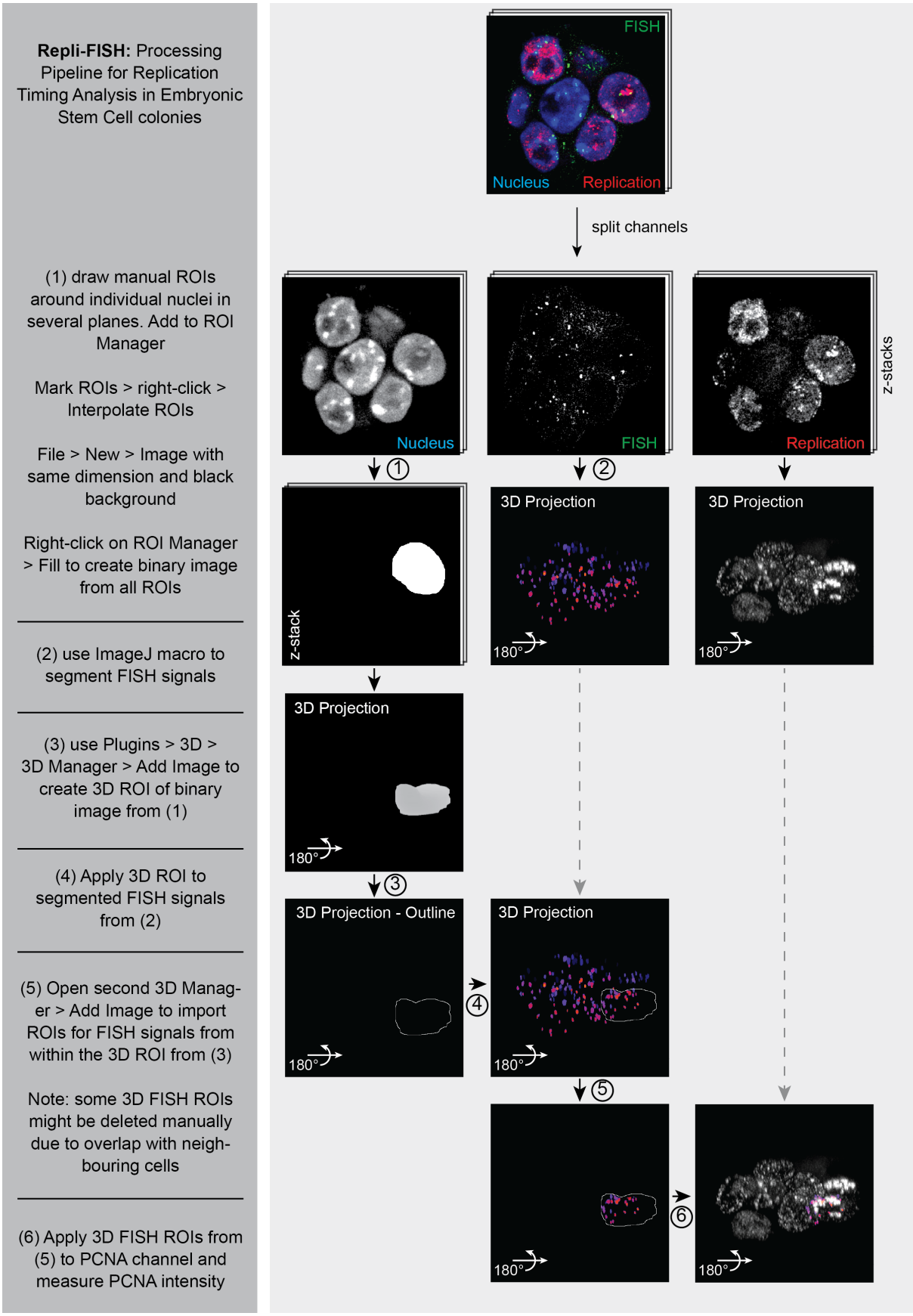


Figure S 4 - Repli-FISH analysis pipeline.
Overview and description of the Repli-FISH analysis pipeline performed for minor satellite and telomere FISH signals

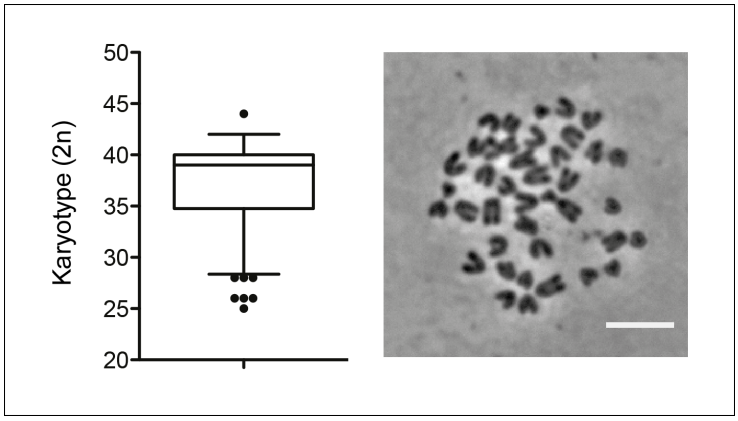


Figure S 5 – Karyotype based genome size estimation. The ploidy of the embryonic stem cell line used for this study was evaluated by karyotype analysis after preparation of metaphase spreads (right, Scale bar = 5 µm). Analysis of > 100 spreads (N = 146) suggests a diploid karyotype (error bars = 5-95 percentiles) corresponding to a genome size of approx. 5.2 Gbp, based on published sizes of individual mouse chromosomes.

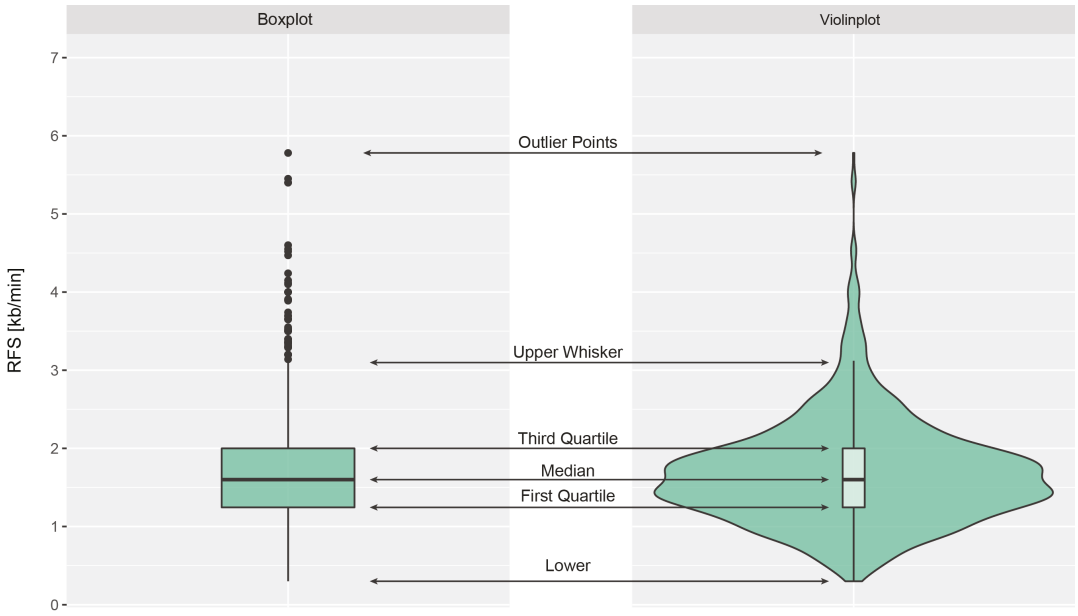


Figure S 6 - Representation of statistical parameters by box- or violin plots

Note: In all tables SD = standard deviation; SEM = standard error of the mean; 95CI = 95% confidence interval

Table S 1 - Manual quantification of tandem repeat elements

Tandem Repeat Element	N_{cells} =	Mean	Median	SD	SEM	95CI
Major Satellite	74	16.68	17	3.0	0.35	0.7
Minor Satellite	54	26.7	26	4.8	0.65	1.29
Telomeres	53	58.43	56	10.0	1.39	2.77

Table S 2 - Summary of 3D-SIM replication foci (BrdU) measurements in embryonic stem cells

S-phase substage	$N_{BrdU-RFi}$ =	Mean	Median	SD	SEM	95CI
I	22	4675.18	4813	1114.21	237.55	494.01
II	25	4879.96	5051	657.82	131.56	271.54
III	13	4806.77	5214	1026.87	284.80	620.53
IV	19	4767.63	4576	1028.05	235.85	495.51
V	3	4411.33	4574	590.06	340.67	1465.80

Table S 3 - Quantification of total PCNA fluorescence intensity within 3D-SIM replication foci (BrdU)

Cell type	$N_{BrdU-RFi}$ =	Mean	Median	SD	SEM	95CI
mouse myoblasts	3399	522152.8	389708	489365.4	8393.783	16457.374
human	3231	294871.2	238579	238700	4199.367	8233.694
mESC	2864	695747	457485.5	761120.9	14222.211	27886.811

Table S 4 - Numbers of nano RFi within pseudoWF RFi

Cell type	N =	Mean	Median	SD	SEM	95CI
mMyoblast	15725	3.1	2	2.33	0.02	0.04
mESC	13489	3.2	2	2.62	0.02	0.04

Table S 5 - Correlation nanoRFi per pseudoWF RFi

Cell type	binned Volume [µm^3]	N =	Mean	Median	SD	SEM	95CI	% of total
mMyoblasts								
	(0,0.2]	13789	2.47	2	1.47	0.01	0.02	87.69%
	(0.2,0.4]	1682	6.69	6	2.36	0.06	0.11	10.70%
	(0.4,0.6]	224	10.75	11	2.88	0.19	0.38	1.42%
	(0.6,0.8]	28	14.21	14	3.18	0.60	1.23	0.18%
	(0.8,1]	2	15.50	15.5	0.71	0.50	6.35	0.01%
	total	15725						
mESC								
	(0,0.2]	11220	2.36	2	1.41	0.01	0.03	83.14%
	(0.2,0.4]	1817	6.18	6	2.38	0.06	0.11	13.46%
	(0.4,0.6]	371	10.30	10	2.96	0.15	0.30	2.75%
	(0.6,0.8]	74	13.74	14	3.87	0.45	0.90	0.55%
	(0.8,1]	7	21.86	23	4.22	1.60	3.90	0.05%
	total	13489						

Table S 6 - Volume measurements [µm^3] of pseudoWF RFi

Cell type	N =	Mean	Median	SD	SEM	95CI
mMyoblast	15725	0.1	0.07	0.09	7	1
mESC	13489	0.12	0.08	0.11	9	2

Cell type	binned Volume [µm^3]	N =	Mean	Median	SD	SEM	95CI	% of total
mMyoblasts								
	(0,0.2]	13789	0.08	0.06	0.05	0.00	0.00	87.69%
	(0.2,0.4]	1682	0.27	0.26	0.05	0.00	0.00	10.70%
	(0.4,0.6]	224	0.47	0.46	0.05	0.00	0.01	1.42%
	(0.6,0.8]	28	0.68	0.67	0.05	0.01	0.02	0.18%
	(0.8,1]	2	0.88	0.88	0.08	0.05	0.68	0.01%
	total	15725						
mESC								
	(0,0.2]	11220	0.08	0	0.05	0.00	0.00	83.14%
	(0.2,0.4]	1817	0.27	0	0.05	0.00	0.00	13.46%
	(0.4,0.6]	371	0.48	0	0.05	0.00	0.01	2.75%
	(0.6,0.8]	74	0.68	1	0.06	0.01	0.01	0.55%
	(0.8,1]	7	0.90	1	0.07	0.03	0.07	0.05%
	total	13489						

Table S 7 - Volume measurements [µm^3] of nano RFi

Cell type	N =	Mean	Median	SD	SEM	95CI
mMyoblast	87805	0.014	0.009	0.02	54	1
mESC	95465	0.014	0.006	0.02	65	1

Cell type	binned Volume [µm^3]	N =	Mean	Median	SD	SEM	95CI	% of total
mMyoblasts								
	(0,0.1]	87598	0.01	0.01	0.02	0.00	0.00	99.76%
	(0.1,0.2]	203	0.12	0.11	0.02	0.00	0.00	0.23%
	(0.2,0.3]	4	0.22	0.22	0.01	0.01	0.02	0.00%
	(0.3,0.4]	0	0	0	0	0	0	0.00%
	total	87805						
mESC								
	(0,0.1]	94735	0.01	0	0.02	0.00	0.00	99.24%
	(0.1,0.2]	700	0.13	0	0.02	0.00	0.00	0.73%
	(0.2,0.3]	28	0.22	0	0.02	0.00	0.01	0.03%
	(0.3,0.4]	2	0.32	0	0.01	0.01	0.13	0.00%
	total	95465						

REFERENCES

Abdurashidova, G., Deganuto, M., Klima, R., Riva, S., Biamonti, G., Giacca, M., and Falaschi, A. (2000). Start sites of bidirectional DNA synthesis at the human lamin B2 origin. Science *287*, 2023-2026.

Abid Ali, F., Douglas, M.E., Locke, J., Pye, V.E., Nans, A., Diffley, J.F.X., and Costa, A. (2017). Cryo-EM structure of a licensed DNA replication origin. Nat Commun *8*, 2241.

Agalioti, T., Chen, G., and Thanos, D. (2002). Deciphering the transcriptional histone acetylation code for a human gene. Cell *111*, 381-392.

Agarwal, N., Becker, A., Jost, K.L., Haase, S., Thakur, B.K., Brero, A., Hardt, T., Kudo, S., Leonhardt, H., and Cardoso, M.C. (2011). MeCP2 Rett mutations affect large scale chromatin organization. Hum Mol Genet *20*, 4187-4195.

Agarwal, N., Hardt, T., Brero, A., Nowak, D., Rothbauer, U., Becker, A., Leonhardt, H., and Cardoso, M.C. (2007). MeCP2 interacts with HP1 and modulates its heterochromatin association during myogenic differentiation. Nucleic Acids Res *35*, 5402-5408.

Ahmad, K., and Henikoff, S. (2001). Centromeres are specialized replication domains in heterochromatin. J Cell Biol *153*, 101-110.

Ahuja, A.K., Jodkowska, K., Teloni, F., Bizard, A.H., Zellweger, R., Herrador, R., Ortega, S., Hickson, I.D., Altmeyer, M., Mendez, J.*, et al.* (2016). A short G1 phase imposes constitutive replication stress and fork remodelling in mouse embryonic stem cells. Nat Commun *7*, 10660.

Aken, B.L., Ayling, S., Barrell, D., Clarke, L., Curwen, V., Fairley, S., Fernandez Banet, J., Billis, K., Garcia Giron, C., Hourlier, T.*, et al.* (2016). The Ensembl gene annotation system. Database (Oxford) *2016*.

Aladjem, M.I. (2007). Replication in context: dynamic regulation of DNA replication patterns in metazoans. Nat Rev Genet *8*, 588-600.

Alberts, B. (2008). Molecular biology of the cell. (New York, NY [u.a.]: Garland Science Taylor & Francis).
Almouzni, G., and Probst, A.V. (2011). Heterochromatin maintenance and establishment: lessons from the mouse pericentromere. Nucleus *2*, 332-338.

Anton, T., Bultmann, S., Leonhardt, H., and Markaki, Y. (2014). Visualization of specific DNA sequences in living mouse embryonic stem cells with a programmable fluorescent CRISPR/Cas system. Nucleus *5*, 163-172.

Arias, E.E., and Walter, J.C. (2005). Replication-dependent destruction of Cdt1 limits DNA replication to a single round per cell cycle in Xenopus egg extracts. Genes Dev *19*, 114-126.

Arita, K., Ariyoshi, M., Tochio, H., Nakamura, Y., and Shirakawa, M. (2008). Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. Nature *455*, 818-821.

Arnoult, N., Schluth-Bolard, C., Letessier, A., Drascovic, I., Bouarich-Bourimi, R., Campisi, J., Kim, S.H., Boussouar, A., Ottaviani, A., Magdinier, F.*, et al.* (2010). Replication timing of human telomeres is chromosome arm-specific, influenced by subtelomeric structures and connected to nuclear localization. PLoS Genet *6*, e1000920.

Aze, A., Sannino, V., Soffientini, P., Bachi, A., and Costanzo, V. (2016). Centromeric DNA replication reconstitution reveals DNA loops and ATR checkpoint suppression. Nat Cell Biol *18*, 684-691.

Ballabeni, A., Park, I.H., Zhao, R., Wang, W., Lerou, P.H., Daley, G.Q., and Kirschner, M.W. (2011). Cell cycle adaptations of embryonic stem cells. Proc Natl Acad Sci U S A *108*, 19252-19257.

Bartke, T., Vermeulen, M., Xhemalce, B., Robson, S.C., Mann, M., and Kouzarides, T. (2010). Nucleosome-interacting proteins regulated by DNA and histone methylation. Cell *143*, 470-484.

Berezney, R., Dubey, D.D., and Huberman, J.A. (2000). Heterogeneity of eukaryotic replicons, replicon clusters, and replication foci. Chromosoma *108*, 471-484.

Bermudez, V.P., Maniwa, Y., Tappin, I., Ozato, K., Yokomori, K., and Hurwitz, J. (2003). The alternative Ctf18-Dcc1-Ctf8-replication factor C complex required for sister chromatid cohesion loads proliferating cell nuclear antigen onto DNA. Proc Natl Acad Sci U S A *100*, 10237-10242.

Besnard, E., Babled, A., Lapasset, L., Milhavet, O., Parrinello, H., Dantec, C., Marin, J.M., and Lemaitre, J.M. (2012). Unraveling cell type-specific and reprogrammable human replication origin signatures associated with G-quadruplex consensus motifs. Nat Struct Mol Biol *19*, 837-844.

Bestor, T., Laudano, A., Mattaliano, R., and Ingram, V. (1988). Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. J Mol Biol *203*, 971-983.

Bialic, M., Coulon, V., Drac, M., Gostan, T., and Schwob, E. (2015). Analyzing the dynamics of DNA replication in Mammalian cells using DNA combing. Methods Mol Biol *1300*, 67-78.

Bianchi, A., and Shore, D. (2007). Early replication of short telomeres in budding yeast. Cell *128*, 1051-1062.

Bianco, J.N., Poli, J., Saksouk, J., Bacal, J., Silva, M.J., Yoshida, K., Lin, Y.L., Tourriere, H., Lengronne, A., and Pasero, P. (2012). Analysis of DNA replication profiles in budding yeast and mammalian cells using DNA combing. Methods *57*, 149-157.

Bickmore, W.A., and Carothers, A.D. (1995). Factors affecting the timing and imprinting of replication on a mammalian chromosome. J Cell Sci *108 (Pt 8)*, 2801-2809.

Bird, A.P., and Wolffe, A.P. (1999). Methylation-induced repression--belts, braces, and chromatin. Cell *99*, 451-454.

Blackburn, E.H. (2001). Switching and signaling at the telomere. Cell *106*, 661-673.

Blow, J.J., Gillespie, P.J., Francis, D., and Jackson, D.A. (2001). Replication origins in Xenopus egg extract Are 5-15 kilobases apart and are activated in clusters that fire at different times. J Cell Biol *152*, 15-25.

Blumenthal, A.B., Kriegstein, H.J., and Hogness, D.S. (1974). The units of DNA replication in Drosophila melanogaster chromosomes. Cold Spring Harb Symp Quant Biol *38*, 205-223.

Boehm, E.M., Gildenberg, M.S., and Washington, M.T. (2016). The Many Roles of PCNA in Eukaryotic DNA Replication. Enzymes *39*, 231-254.

Bogan, J.A., Natale, D.A., and Depamphilis, M.L. (2000). Initiation of eukaryotic DNA replication: conservative or liberal? J Cell Physiol *184*, 139-150.

Bousset, K., and Diffley, J.F. (1998). The Cdc7 protein kinase is required for origin firing during S phase. Genes Dev *12*, 480-490.

Braunstein, J.D., Schulze, D., DelGiudice, T., Furst, A., and Schildkraut, C.L. (1982). The temporal order of replication of murine immunoglobulin heavy chain constant region sequences corresponds to their linear order in the genome. Nucleic Acids Res *10*, 6887-6902.

Brero, A., Easwaran, H.P., Nowak, D., Grunewald, I., Cremer, T., Leonhardt, H., and Cardoso, M.C. (2005). Methyl CpG-binding proteins induce large-scale chromatin reorganization during terminal differentiation. J Cell Biol *169*, 733-743.

Buongiorno-Nardelli, M., Micheli, G., Carri, M.T., and Marilley, M. (1982). A relationship between replicon size and supercoiled loop domains in the eukaryotic genome. Nature *298*, 100-102.

Burdon, T., Smith, A., and Savatier, P. (2002). Signalling, cell cycle and pluripotency in embryonic stem cells. Trends Cell Biol *12*, 432-438.

Burhans, W.C., Vassilev, L.T., Caddle, M.S., Heintz, N.H., and DePamphilis, M.L. (1990). Identification of an origin of bidirectional DNA replication in mammalian chromosomes. Cell *62*, 955-965.

Cairns, J. (1966). Autoradiography of HeLa cell DNA. J Mol Biol *15*, 372-373.

Callan, H.G. (1974). DNA replication in the chromosomes of eukaryotes. Cold Spring Harb Symp Quant Biol *38*, 195-203.

Cardoso, M.C., Schneider, K., Martin, R.M., and Leonhardt, H. (2012). Structure, function and dynamics of nuclear subcompartments. Curr Opin Cell Biol *24*, 79-85.

Casas-Delucchi, C.S., Brero, A., Rahn, H.P., Solovei, I., Wutz, A., Cremer, T., Leonhardt, H., and Cardoso, M.C. (2011). Histone acetylation controls the inactive X chromosome replication dynamics. Nat Commun *2*, 222.

Casas-Delucchi, C.S., and Cardoso, M.C. (2011). Epigenetic control of DNA replication dynamics in mammals. Nucleus *2*, 370-382.

Casas-Delucchi, C.S., van Bemmel, J.G., Haase, S., Herce, H.D., Nowak, D., Meilinger, D., Stear, J.H., Leonhardt, H., and Cardoso, M.C. (2012). Histone hypoacetylation is required to maintain late replication timing of constitutive heterochromatin. Nucleic Acids Res *40*, 159-169.

Celeda, D., Aldinger, K., Haar, F.M., Hausmann, M., Durm, M., Ludwig, H., and Cremer, C. (1994). Rapid fluorescence in situ hybridization with repetitive DNA probes: quantification by digital image analysis. Cytometry *17*, 13-25.

Celeda, D., Bettag, U., and Cremer, C. (1992). A simplified combination of DNA probe preparation and fluorescence in situ hybridization. Z Naturforsch C *47*, 739-747.

Chagin, V.O., Casas-Delucchi, C.S., Reinhart, M., Schermelleh, L., Markaki, Y., Maiser, A., Bolius, J.J., Bensimon, A., Fillies, M., Domaing, P.*, et al.* (2016). 4D Visualization of replication foci in mammalian cells corresponding to individual replicons. Nat Commun *7*, 11231.

Chagin, V.O., Reinhart, M., and Cardoso, M.C. (2015). High-resolution analysis of Mammalian DNA replication units. Methods Mol Biol *1300*, 43-65.

Chagin, V.O., Stear, J.H., and Cardoso, M.C. (2010). Organization of DNA replication. Cold Spring Harb Perspect Biol *2*, a000737.

Chen, T., Ueda, Y., Dodge, J.E., Wang, Z., and Li, E. (2003). Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. Mol Cell Biol *23*, 5594-5605.

Chen, W.Y., and Townes, T.M. (2000). Molecular mechanism for silencing virally transduced genes involves histone deacetylation and chromatin condensation. Proc Natl Acad Sci U S A *97*, 377-382.

Cheng, M.K., and Disteche, C.M. (2004). Silence of the fathers: early X inactivation. Bioessays *26*, 821-824.

Chilkova, O., Stenlund, P., Isoz, I., Stith, C.M., Grabowski, P., Lundstrom, E.B., Burgers, P.M., and Johansson, E. (2007). The eukaryotic leading and lagging strand DNA polymerases are loaded onto primer-ends via separate mechanisms but have comparable processivity in the presence of PCNA. Nucleic Acids Res *35*, 6588-6597.

Chuang, L.S., Ian, H.I., Koh, T.W., Ng, H.H., Xu, G., and Li, B.F. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. Science *277*, 1996-2000.

Clouaire, T., and Stancheva, I. (2008). Methyl-CpG binding proteins: specialized transcriptional repressors or structural components of chromatin? Cell Mol Life Sci *65*, 1509-1522.

Collins, N., Poot, R.A., Kukimoto, I., Garcia-Jimenez, C., Dellaire, G., and Varga-Weisz, P.D. (2002). An ACF1-ISWI chromatin-remodeling complex is required for DNA replication through heterochromatin. Nat Genet *32*, 627-632.

Conti, C., Caburet, S., Schurra, C., and Bensimon, A. (2001). Molecular combing. Curr Protoc Cytom *Chapter 8*, Unit 8 10.

Conti, C., Sacca, B., Herrick, J., Lalou, C., Pommier, Y., and Bensimon, A. (2007). Replication fork velocities at adjacent replication origins are coordinately modified during DNA replication in human cells. Mol Biol Cell *18*, 3059-3067.

Coster, G., and Diffley, J.F.X. (2017). Bidirectional eukaryotic DNA replication is established by quasi-symmetrical helicase loading. Science *357*, 314-318.

Coster, G., Frigola, J., Beuron, F., Morris, E.P., and Diffley, J.F. (2014). Origin licensing requires ATP binding and hydrolysis by the MCM replicative helicase. Mol Cell *55*, 666-677.

Courbet, S., Gay, S., Arnoult, N., Wronka, G., Anglana, M., Brison, O., and Debatisse, M. (2008). Replication fork movement sets chromatin loop size and origin choice in mammalian cells. Nature *455*, 557-560.

Crabbe, L., Verdun, R.E., Haggblom, C.I., and Karlseder, J. (2004). Defective telomere lagging strand synthesis in cells lacking WRN helicase activity. Science *306*, 1951-1953.

Cremer, M., Grasser, F., Lanctot, C., Muller, S., Neusser, M., Zinner, R., Solovei, I., and Cremer, T. (2008). Multicolor 3D fluorescence in situ hybridization for imaging interphase chromosomes. Methods Mol Biol *463*, 205-239.

Cremer, T., and Cremer, C. (2001). Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet *2*, 292-301.

Cremer, T., and Cremer, M. (2010). Chromosome territories. Cold Spring Harb Perspect Biol *2*, a003889.

Csink, A.K., and Henikoff, S. (1998). Something from nothing: the evolution and utility of satellite repeats. Trends Genet *14*, 200-204.

Darzynkiewicz, Z., Huang, X., and Zhao, H. (2017). Analysis of Cellular DNA Content by Flow Cytometry. Curr Protoc Cytom *82*, 7 5 1-7 5 20.

Dawlaty, M.M., Breiling, A., Le, T., Raddatz, G., Barrasa, M.I., Cheng, A.W., Gao, Q., Powell, B.E., Li, Z., Xu, M.*, et al.* (2013). Combined deficiency of Tet1 and Tet2 causes epigenetic abnormalities but is compatible with postnatal development. Dev Cell *24*, 310-323.

de Koning, A.P., Gu, W., Castoe, T.A., Batzer, M.A., and Pollock, D.D. (2011). Repetitive elements may comprise over two-thirds of the human genome. PLoS Genet *7*, e1002384.

de Lange, T. (2005). Shelterin: the protein complex that shapes and safeguards human telomeres. Genes Dev *19*, 2100-2110.

De March, M., Merino, N., Barrera-Vilarmau, S., Crehuet, R., Onesti, S., Blanco, F.J., and De Biasio, A. (2017). Structural basis of human PCNA sliding on DNA. Nat Commun *8*, 13935.

Deegan, T.D., Yeeles, J.T., and Diffley, J.F. (2016). Phosphopeptide binding by Sld3 links Dbf4-dependent kinase to MCM replicative helicase activation. EMBO J *35*, 961-973.

Dekker, J., and Mirny, L. (2016). The 3D Genome as Moderator of Chromosomal Communication. Cell *164*, 1110-1121.

Deng, X., Berletch, J.B., Nguyen, D.K., and Disteche, C.M. (2014). X chromosome regulation: diverse patterns in development, tissues and disease. Nat Rev Genet *15*, 367-378.

Deng, X., Ma, W., Ramani, V., Hill, A., Yang, F., Ay, F., Berletch, J.B., Blau, C.A., Shendure, J., Duan, Z.*, et al.* (2015). Bipartite structure of the inactive mouse X chromosome. Genome Biol *16*, 152.

Devbhandari, S., Jiang, J., Kumar, C., Whitehouse, I., and Remus, D. (2017). Chromatin Constrains the Initiation and Elongation of DNA Replication. Mol Cell *65*, 131-141.

Dijkwel, P.A., and Hamlin, J.L. (1995). The Chinese hamster dihydrofolate reductase origin consists of multiple potential nascent-strand start sites. Mol Cell Biol *15*, 3023-3031.

Dimitrova, D.S., and Gilbert, D.M. (1999). The spatial position and replication timing of chromosomal domains are both established in early G1 phase. Mol Cell *4*, 983-993.

Dixon, J.R., Jung, I., Selvaraj, S., Shen, Y., Antosiewicz-Bourget, J.E., Lee, A.Y., Ye, Z., Kim, A., Rajagopal, N., Xie, W.*, et al.* (2015). Chromatin architecture reorganization during stem cell differentiation. Nature *518*, 331-336.

Dixon, J.R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., Hu, M., Liu, J.S., and Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. Nature *485*, 376-380.

Doetschman, T., Gregg, R.G., Maeda, N., Hooper, M.L., Melton, D.W., Thompson, S., and Smithies, O. (1987). Targetted correction of a mutant HPRT gene in mouse embryonic stem cells. Nature *330*, 576-578.

Donaldson, A.D. (2005). Shaping time: chromatin structure and the DNA replication programme. Trends Genet *21*, 444-449.

Douglas, M.E., Ali, F.A., Costa, A., and Diffley, J.F.X. (2018). The mechanism of eukaryotic CMG helicase activation. Nature *555*, 265-268.

Douglas, M.E., and Diffley, J.F. (2012). Replication timing: the early bird catches the worm. Curr Biol *22*, R81-82.

Dowell, S.J., Romanowski, P., and Diffley, J.F. (1994). Interaction of Dbf4, the Cdc7 protein kinase regulatory subunit, with yeast replication origins in vivo. Science *265*, 1243-1246.

Easwaran, H.P., Schermelleh, L., Leonhardt, H., and Cardoso, M.C. (2004). Replication-independent chromatin loading of Dnmt1 during G2 and M phases. EMBO Rep *5*, 1181-1186.

Edenberg, H.J., and Huberman, J.A. (1975). Eukaryotic chromosome replication. Annu Rev Genet *9*, 245-284.

Edgar, B.A., and Schubiger, G. (1986). Parameters controlling transcriptional activation during early Drosophila development. Cell *44*, 871-877.

Edwards, M.C., Tutter, A.V., Cvetic, C., Gilbert, C.H., Prokhorova, T.A., and Walter, J.C. (2002). MCM2-7 complexes bind chromatin in a distributed pattern surrounding the origin recognition complex in Xenopus egg extracts. J Biol Chem *277*, 33049-33057.

Efroni, S., Duttagupta, R., Cheng, J., Dehghani, H., Hoeppner, D.J., Dash, C., Bazett-Jones, D.P., Le Grice, S., McKay, R.D., Buetow, K.H.*, et al.* (2008). Global transcription in pluripotent embryonic stem cells. Cell Stem Cell *2*, 437-447.

Eissenberg, J.C., and Elgin, S.C. (2000). The HP1 protein family: getting a grip on chromatin. Curr Opin Genet Dev *10*, 204-210.

Enemark, E.J., and Joshua-Tor, L. (2008). On helicases and other motor proteins. Curr Opin Struct Biol *18*, 243-257.

Ermakova, O.V., Nguyen, L.H., Little, R.D., Chevillard, C., Riblet, R., Ashouian, N., Birshtein, B.K., and Schildkraut, C.L. (1999). Evidence that a single replication fork proceeds from early to late replicating domains in the IgH locus in a non-B cell line. Mol Cell *3*, 321-330.

Esteve, P.O., Chin, H.G., Smallwood, A., Feehery, G.R., Gangisetty, O., Karpf, A.R., Carey, M.F., and Pradhan, S. (2006). Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. Genes Dev *20*, 3089-3103.

Evrin, C., Clarke, P., Zech, J., Lurz, R., Sun, J., Uhle, S., Li, H., Stillman, B., and Speck, C. (2009). A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. Proc Natl Acad Sci U S A *106*, 20240-20245.

Felsenfeld, G., and Groudine, M. (2003). Controlling the double helix. Nature *421*, 448-453.

Ferreira, J., and Carmo-Fonseca, M. (1997). Genome replication in early mouse embryos follows a defined temporal and spatial order. J Cell Sci *110 (Pt 7)*, 889-897.

Ferrell, J.E., Jr., Wu, M., Gerhart, J.C., and Martin, G.S. (1991). Cell cycle tyrosine phosphorylation of p34cdc2 and a microtubule-associated protein kinase homolog in Xenopus oocytes and eggs. Mol Cell Biol *11*, 1965-1971.

Filippova, G.N., Fagerlie, S., Klenova, E.M., Myers, C., Dehner, Y., Goodwin, G., Neiman, P.E., Collins, S.J., and Lobanenkov, V.V. (1996). An exceptionally conserved transcriptional repressor, CTCF, employs different combinations of zinc fingers to bind diverged promoter sequences of avian and mammalian c-myc oncogenes. Mol Cell Biol *16*, 2802-2813.

Flockinger, R.A., Freedman, M.L., and Stambrook, P.J. (1967). Generation times and DNA replication patterns of cells of developing frog embryos. Dev Biol *16*, 457-473.

Fragkos, M., Ganier, O., Coulombe, P., and Mechali, M. (2015). DNA replication origin activation in space and time. Nat Rev Mol Cell Biol *16*, 360-374.

Fraser, P. (2006). Transcriptional control thrown for a loop. Curr Opin Genet Dev *16*, 490-495.

Frauer, C., Rottach, A., Meilinger, D., Bultmann, S., Fellinger, K., Hasenoder, S., Wang, M., Qin, W., Soding, J., Spada, F.*, et al.* (2011). Different binding properties and function of CXXC zinc finger domains in Dnmt1 and Tet1. PLoS One *6*, e16627.

Friedman, K.L., Brewer, B.J., and Fangman, W.L. (1997). Replication profile of Saccharomyces cerevisiae chromosome VI. Genes Cells *2*, 667-678.

Friedman, K.L., Raghuraman, M.K., Fangman, W.L., and Brewer, B.J. (1995). Analysis of the temporal program of replication initiation in yeast chromosomes. J Cell Sci Suppl *19*, 51-58.

Fu, Y.V., Yardimci, H., Long, D.T., Ho, T.V., Guainazzi, A., Bermudez, V.P., Hurwitz, J., van Oijen, A., Scharer, O.D., and Walter, J.C. (2011). Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. Cell *146*, 931-941.

Fudenberg, G., Imakaev, M., Lu, C., Goloborodko, A., Abdennur, N., and Mirny, L.A. (2016). Formation of Chromosomal Domains by Loop Extrusion. Cell Rep *15*, 2038-2049.

Fujii-Yamamoto, H., Kim, J.M., Arai, K., and Masai, H. (2005). Cell cycle and developmental regulations of replication factors in mouse embryonic stem cells. J Biol Chem *280*, 12976-12987.

Gaspar-Maia, A., Alajem, A., Polesso, F., Sridharan, R., Mason, M.J., Heidersbach, A., Ramalho-Santos, J., McManus, M.T., Plath, K., Meshorer, E.*, et al.* (2009). Chd1 regulates open chromatin and pluripotency of embryonic stem cells. Nature *460*, 863-868.

Gavin, K.A., Hidaka, M., and Stillman, B. (1995). Conserved initiator proteins in eukaryotes. Science *270*, 1667-1671.

Ge, X.Q., Han, J., Cheng, E.C., Yamaguchi, S., Shima, N., Thomas, J.L., and Lin, H. (2015). Embryonic Stem Cells License a High Level of Dormant Origins to Protect the Genome against Replication Stress. Stem Cell Reports *5*, 185-194.

Ge, X.Q., Jackson, D.A., and Blow, J.J. (2007). Dormant origins licensed by excess Mcm2-7 are required for human cells to survive replicative stress. Genes Dev *21*, 3331-3341.

Georgescu, R., Yuan, Z., Bai, L., de Luna Almeida Santos, R., Sun, J., Zhang, D., Yurieva, O., Li, H., and O'Donnell, M.E. (2017). Structure of eukaryotic CMG helicase at a replication fork and implications to replisome architecture and origin initiation. Proc Natl Acad Sci U S A *114*, E697-E706.

Ghosh, M., Kemp, M., Liu, G., Ritzi, M., Schepers, A., and Leffak, M. (2006). Differential binding of replication proteins across the human c-myc replicator. Mol Cell Biol *26*, 5270-5283.

Gilson, E., and Geli, V. (2007). How telomeres are replicated. Nat Rev Mol Cell Biol *8*, 825-838.

Glynn, E.F., Megee, P.C., Yu, H.G., Mistrot, C., Unal, E., Koshland, D.E., DeRisi, J.L., and Gerton, J.L. (2004). Genome-wide mapping of the cohesin complex in the yeast Saccharomyces cerevisiae. PLoS Biol *2*, E259.

Goldar, A., Labit, H., Marheineke, K., and Hyrien, O. (2008). A dynamic stochastic model for DNA replication initiation in early embryos. PLoS One *3*, e2919.

Gopalakrishnan, S., Van Emburgh, B.O., and Robertson, K.D. (2008). DNA methylation in development and human disease. Mutat Res *647*, 30-38.

Gossen, M., Pak, D.T., Hansen, S.K., Acharya, J.K., and Botchan, M.R. (1995). A Drosophila homolog of the yeast origin recognition complex. Science *270*, 1674-1677.

Griffith, J.D., Comeau, L., Rosenfield, S., Stansel, R.M., Bianchi, A., Moss, H., and de Lange, T. (1999). Mammalian telomeres end in a large duplex loop. Cell *97*, 503-514.

Guacci, V., Koshland, D., and Strunnikov, A. (1997). A direct link between sister chromatid cohesion and chromosome condensation revealed through the analysis of MCD1 in S. cerevisiae. Cell *91*, 47-57.

Guenatri, M., Bailly, D., Maison, C., and Almouzni, G. (2004). Mouse centric and pericentric satellite repeats form distinct functional heterochromatin. J Cell Biol *166*, 493-505.

Guillou, E., Ibarra, A., Coulon, V., Casado-Vela, J., Rico, D., Casal, I., Schwob, E., Losada, A., and Mendez, J. (2010). Cohesin organizes chromatin loops at DNA replication factories. Genes Dev *24*, 2812-2822.

Gullerova, M., and Proudfoot, N.J. (2008). Cohesin complex promotes transcriptional termination between convergent genes in S. pombe. Cell *132*, 983-995.

Haaf, T., and Schmid, M. (2000). Experimental condensation inhibition in constitutive and facultative heterochromatin of mammalian chromosomes. Cytogenet Cell Genet *91*, 113-123.

Hadjur, S., Williams, L.M., Ryan, N.K., Cobb, B.S., Sexton, T., Fraser, P., Fisher, A.G., and Merkenschlager, M. (2009). Cohesins form chromosomal cis-interactions at the developmentally regulated IFNG locus. Nature *460*, 410-413.

Hahn, A.T., Jones, J.T., and Meyer, T. (2009). Quantitative analysis of cell cycle phase durations and PC12 differentiation using fluorescent biosensors. Cell Cycle *8*, 1044-1052.

Han, J., Zhou, H., Horazdovsky, B., Zhang, K., Xu, R.M., and Zhang, Z. (2007). Rtt109 acetylates histone H3 lysine 56 and functions in DNA replication. Science *315*, 653-655.

Handoko, L., Xu, H., Li, G., Ngan, C.Y., Chew, E., Schnapp, M., Lee, C.W., Ye, C., Ping, J.L., Mulawadi, F.*, et al.* (2011). CTCF-mediated functional chromatin interactome in pluripotent cells. Nat Genet *43*, 630-638.

Hargreaves, D.C., and Crabtree, G.R. (2011). ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. Cell Res *21*, 396-420.

Harvey, K.J., and Newport, J. (2003). Metazoan origin selection: origin recognition complex chromatin binding is regulated by CDC6 recruitment and ATP hydrolysis. J Biol Chem *278*, 48524-48528.

Hayashi, M.T., Takahashi, T.S., Nakagawa, T., Nakayama, J., and Masukata, H. (2009). The heterochromatin protein Swi6/HP1 activates replication origins at the pericentromeric region and silent mating-type locus. Nat Cell Biol *11*, 357-362.

Heinz, K.S., Casas-Delucchi, C.S., Torok, T., Cmarko, D., Rapp, A., Raska, I., and Cardoso, M.C. (2018). Peripheral re-localization of constitutive heterochromatin advances its replication timing and impairs maintenance of silencing marks. Nucleic Acids Res *46*, 6112-6128.

Hermann, A., Goyal, R., and Jeltsch, A. (2004). The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. J Biol Chem *279*, 48350-48359.

Hiratani, I., Ryba, T., Itoh, M., Rathjen, J., Kulik, M., Papp, B., Fussner, E., Bazett-Jones, D.P., Plath, K., Dalton, S.*, et al.* (2010). Genome-wide dynamics of replication timing revealed by in vitro models of mouse embryogenesis. Genome Res *20*, 155-169.

Hiratani, I., Ryba, T., Itoh, M., Yokochi, T., Schwaiger, M., Chang, C.W., Lyou, Y., Townes, T.M., Schubeler, D., and Gilbert, D.M. (2008). Global reorganization of replication domains during embryonic stem cell differentiation. PLoS Biol *6*, e245.

Hiyama, E., and Hiyama, K. (2007). Telomere and telomerase in stem cells. Br J Cancer *96*, 1020-1024.

Hotchkiss, R.D. (1948). The quantitative separation of purines, pyrimidines, and nucleosides by paper chromatography. J Biol Chem *175*, 315-332.

Hou, C., Dale, R., and Dean, A. (2010). Cell type specificity of chromatin organization mediated by CTCF and cohesin. Proc Natl Acad Sci U S A *107*, 3651-3656.

Houlard, M., Berlivet, S., Probst, A.V., Quivy, J.P., Hery, P., Almouzni, G., and Gerard, M. (2006). CAF-1 is essential for heterochromatin organization in pluripotent embryonic cells. PLoS Genet *2*, e181.

Hozak, P., Jackson, D.A., and Cook, P.R. (1994). Replication factories and nuclear bodies: the ultrastructural characterization of replication sites during the cell cycle. J Cell Sci *107 (Pt 8)*, 2191-2202.

Huberman, J.A., and Riggs, A.D. (1968). On the mechanism of DNA replication in mammalian chromosomes. J Mol Biol *32*, 327-341.

Hubner, M.R., Eckersley-Maslin, M.A., and Spector, D.L. (2013). Chromatin organization and transcriptional regulation. Curr Opin Genet Dev *23*, 89-95.

Hyrien, O., Marheineke, K., and Goldar, A. (2003). Paradoxes of eukaryotic DNA replication: MCM proteins and the random completion problem. Bioessays *25*, 116-125.

Hyrien, O., Maric, C., and Mechali, M. (1995). Transition in specification of embryonic metazoan DNA replication origins. Science *270*, 994-997.

Hyrien, O., and Mechali, M. (1993). Chromosomal replication initiates and terminates at random sequences but at regular intervals in the ribosomal DNA of Xenopus early embryos. EMBO J *12*, 4511-4520.

Jackson, D.A., and Pombo, A. (1998). Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. J Cell Biol *140*, 1285-1295.

Jackson, M., Krassowska, A., Gilbert, N., Chevassut, T., Forrester, L., Ansell, J., and Ramsahoye, B. (2004). Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. Mol Cell Biol *24*, 8862-8871.

Jackson-Grusby, L., Beard, C., Possemato, R., Tudor, M., Fambrough, D., Csankovszki, G., Dausman, J., Lee, P., Wilson, C., Lander, E.*, et al.* (2001). Loss of genomic methylation causes p53-dependent apoptosis and epigenetic deregulation. Nat Genet *27*, 31-39.

Jacob, F., and Brenner, S. (1963). [On the regulation of DNA synthesis in bacteria: the hypothesis of the replicon]. C R Hebd Seances Acad Sci *256*, 298-300.

Jeppesen, P., Mitchell, A., Turner, B., and Perry, P. (1992). Antibodies to defined histone epitopes reveal variations in chromatin conformation and underacetylation of centric heterochromatin in human metaphase chromosomes. Chromosoma *101*, 322-332.

Jobling, M.A., and Tyler-Smith, C. (2003). The human Y chromosome: an evolutionary marker comes of age. Nat Rev Genet *4*, 598-612.

Jones, P.A., and Baylin, S.B. (2002). The fundamental role of epigenetic events in cancer. Nat Rev Genet *3*, 415-428.

Jones, P.L., Veenstra, G.J., Wade, P.A., Vermaak, D., Kass, S.U., Landsberger, N., Strouboulis, J., and Wolffe, A.P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat Genet *19*, 187-191.

Jorgensen, H.F., Azuara, V., Amoils, S., Spivakov, M., Terry, A., Nesterova, T., Cobb, B.S., Ramsahoye, B., Merkenschlager, M., and Fisher, A.G. (2007). The impact of chromatin modifiers on the timing of locus replication in mouse embryonic stem cells. Genome Biol *8*, R169.

Kanemaki, M., and Labib, K. (2006). Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks. EMBO J *25*, 1753-1763.

Kang, S., Warner, M.D., and Bell, S.P. (2014). Multiple functions for Mcm2-7 ATPase motifs during replication initiation. Mol Cell *55*, 655-665.

Kim, J.K., Esteve, P.O., Jacobsen, S.E., and Pradhan, S. (2009). UHRF1 binds G9a and participates in p21 transcriptional regulation in mammalian cells. Nucleic Acids Res *37*, 493-505.

Kim, S.M., Dubey, D.D., and Huberman, J.A. (2003). Early-replicating heterochromatin. Genes Dev *17*, 330-335.

Kim, S.M., and Huberman, J.A. (2001). Regulation of replication timing in fission yeast. EMBO J *20*, 6115-6126.

Kimura, H., Morii, E., Ikeda, J.I., Ezoe, S., Xu, J.X., Nakamichi, N., Tomita, Y., Shibayama, H., Kanakura, Y., and Aozasa, K. (2006). Role of DNA methylation for expression of novel stem cell marker CDCP1 in hematopoietic cells. Leukemia *20*, 1551-1556.

Kipling, D., Ackford, H.E., Taylor, B.A., and Cooke, H.J. (1991). Mouse minor satellite DNA genetically maps to the centromere and is physically linked to the proximal telomere. Genomics *11*, 235-241.

Kitamura, E., Blow, J.J., and Tanaka, T.U. (2006). Live-cell imaging reveals replication of individual replicons in eukaryotic replication factories. Cell *125*, 1297-1308.

Klose, R.J., and Bird, A.P. (2006). Genomic DNA methylation: the mark and its mediators. Trends Biochem Sci *31*, 89-97.

Kobayakawa, S., Miike, K., Nakao, M., and Abe, K. (2007). Dynamic changes in the epigenomic state and nuclear organization of differentiating mouse embryonic stem cells. Genes Cells *12*, 447-460.

Koren, A. (2014). DNA replication timing: Coordinating genome stability with genome regulation on the X chromosome and beyond. Bioessays *36*, 997-1004.

Kraus, F., Miron, E., Demmerle, J., Chitiashvili, T., Budco, A., Alle, Q., Matsuda, A., Leonhardt, H., Schermelleh, L., and Markaki, Y. (2017). Quantitative 3D structured illumination microscopy of nuclear structures. Nat Protoc *12*, 1011-1028.

Krivega, I., and Dean, A. (2012). Enhancer and promoter interactions-long distance calls. Curr Opin Genet Dev *22*, 79-85.

Kurat, C.F., Yeeles, J.T.P., Patel, H., Early, A., and Diffley, J.F.X. (2017). Chromatin Controls DNA Replication Origin Selection, Lagging-Strand Synthesis, and Replication Fork Rates. Mol Cell *65*, 117-130.

Kurisaki, A., Hamazaki, T.S., Okabayashi, K., Iida, T., Nishine, T., Chonan, R., Kido, H., Tsunasawa, S., Nishimura, O., Asashima, M.*, et al.* (2005). Chromatin-related proteins in pluripotent mouse embryonic stem cells are downregulated after removal of leukemia inhibitory factor. Biochem Biophys Res Commun *335*, 667-675.

Labib, K., and Gambus, A. (2007). A key role for the GINS complex at DNA replication forks. Trends Cell Biol *17*, 271-278.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116-120.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W.*, et al.* (2001). Initial sequencing and analysis of the human genome. Nature *409*, 860-921.

Lando, D., Stevens, T.J., Basu, S., and Laue, E.D. (2018). Calculation of 3D genome structures for comparison of chromosome conformation capture experiments with microscopy: An evaluation of single-cell Hi-C protocols. Nucleus *9*, 190-201.

Langley, A.R., Graf, S., Smith, J.C., and Krude, T. (2016). Genome-wide identification and characterisation of human DNA replication origins by initiation site sequencing (ini-seq). Nucleic Acids Res *44*, 10230-10247.

Lebofsky, R., and Bensimon, A. (2003). Single DNA molecule analysis: applications of molecular combing. Brief Funct Genomic Proteomic *1*, 385-396.

Lehnertz, B., Ueda, Y., Derijck, A.A., Braunschweig, U., Perez-Burgos, L., Kubicek, S., Chen, T., Li, E., Jenuwein, T., and Peters, A.H. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. Curr Biol *13*, 1192-1200.

Lei, H., Oh, S.P., Okano, M., Juttermann, R., Goss, K.A., Jaenisch, R., and Li, E. (1996). De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. Development *122*, 3195-3205.

Lengronne, A., Katou, Y., Mori, S., Yokobayashi, S., Kelly, G.P., Itoh, T., Watanabe, Y., Shirahige, K., and Uhlmann, F. (2004). Cohesin relocation from sites of chromosomal loading to places of convergent transcription. Nature *430*, 573-578.

Leonhardt, H., Page, A.W., Weier, H.U., and Bestor, T.H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell *71*, 865-873.

Leonhardt, H., Rahn, H.P., Weinzierl, P., Sporbert, A., Cremer, T., Zink, D., and Cardoso, M.C. (2000). Dynamics of DNA replication factories in living cells. J Cell Biol *149*, 271-280.

Levy, M.Z., Allsopp, R.C., Futcher, A.B., Greider, C.W., and Harley, C.B. (1992). Telomere end-replication problem and cell aging. J Mol Biol *225*, 951-960.

Li, B., Su, T., Ferrari, R., Li, J.Y., and Kurdistani, S.K. (2014). A unique epigenetic signature is associated with active DNA replication loci in human embryonic stem cells. Epigenetics *9*, 257-267.

Li, E. (2002). Chromatin modification and epigenetic reprogramming in mammalian development. Nat Rev Genet *3*, 662-673.

Li, E., Bestor, T.H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915-926.

Li, H., and O'Donnell, M.E. (2018). The Eukaryotic CMG Helicase at the Replication Fork: Emerging Architecture Reveals an Unexpected Mechanism. Bioessays *40*.

Li, J., Santoro, R., Koberna, K., and Grummt, I. (2005). The chromatin remodeling complex NoRC controls replication timing of rRNA genes. EMBO J *24*, 120-127.

Li, V.C., Ballabeni, A., and Kirschner, M.W. (2012). Gap 1 phase length and mouse embryonic stem cell self-renewal. Proc Natl Acad Sci U S A *109*, 12550-12555.

Liang, C., Weinreich, M., and Stillman, B. (1995). ORC and Cdc6p interact and determine the frequency of initiation of DNA replication in the genome. Cell *81*, 667-676.

Liao, J., Karnik, R., Gu, H., Ziller, M.J., Clement, K., Tsankov, A.M., Akopian, V., Gifford, C.A., Donaghey, J., Galonska, C.*, et al.* (2015). Targeted disruption of DNMT1, DNMT3A and DNMT3B in human embryonic stem cells. Nat Genet *47*, 469-478.

Lindhout, B.I., Fransz, P., Tessadori, F., Meckel, T., Hooykaas, P.J., and van der Zaal, B.J. (2007). Live cell imaging of repetitive DNA sequences via GFP-tagged polydactyl zinc finger proteins. Nucleic Acids Res *35*, e107.

Liu, Y., Oakeley, E.J., Sun, L., and Jost, J.P. (1998). Multiple domains are involved in the targeting of the mouse DNA methyltransferase to the DNA replication foci. Nucleic Acids Res *26*, 1038-1045.

Lob, D., Lengert, N., Chagin, V.O., Reinhart, M., Casas-Delucchi, C.S., Cardoso, M.C., and Drossel, B. (2016). 3D replicon distributions arise from stochastic initiation and domino-like DNA replication progression. Nat Commun *7*, 11207.

Lopes Novo, C., and Rugg-Gunn, P.J. (2016). Chromatin organization in pluripotent cells: emerging approaches to study and disrupt function. Brief Funct Genomics *15*, 305-314.

Lopez, C.M., Lloyd, A.J., Leonard, K., and Wilkinson, M.J. (2012). Differential effect of three base modifications on DNA thermostability revealed by high resolution melting. Anal Chem *84*, 7336-7342.

Lukas, C., Savic, V., Bekker-Jensen, S., Doil, C., Neumann, B., Pedersen, R.S., Grofte, M., Chan, K.L., Hickson, I.D., Bartek, J.*, et al.* (2011). 53BP1 nuclear bodies form around DNA lesions generated by mitotic transmission of chromosomes under replication stress. Nat Cell Biol *13*, 243-253.

Ma, H., Samarabandu, J., Devdhar, R.S., Acharya, R., Cheng, P.C., Meng, C., and Berezney, R. (1998). Spatial and temporal dynamics of DNA replication sites in mammalian cells. J Cell Biol *143*, 1415-1425.

MacAlpine, H.K., Gordan, R., Powell, S.K., Hartemink, A.J., and MacAlpine, D.M. (2010). Drosophila ORC localizes to open chromatin and marks sites of cohesin complex loading. Genome Res *20*, 201-211.

Machida, Y.J., Hamlin, J.L., and Dutta, A. (2005). Right place, right time, and only once: replication initiation in metazoans. Cell *123*, 13-24.

Maeshima, K., Hihara, S., and Eltsov, M. (2010). Chromatin structure: does the 30-nm fibre exist in vivo? Curr Opin Cell Biol *22*, 291-297.

Manders, E.M., Stap, J., Brakenhoff, G.J., van Driel, R., and Aten, J.A. (1992). Dynamics of three-dimensional replication patterns during the S-phase, analysed by double labelling of DNA and confocal microscopy. J Cell Sci *103 (Pt 3)*, 857-862.

Manders, E.M., Stap, J., Strackee, J., van Driel, R., and Aten, J.A. (1996). Dynamic behavior of DNA replication domains. Exp Cell Res *226*, 328-335.

Mantiero, D., Mackenzie, A., Donaldson, A., and Zegerman, P. (2011). Limiting replication initiation factors execute the temporal programme of origin firing in budding yeast. EMBO J *30*, 4805-4814.

Maric, C., Benard, M., and Pierron, G. (2003). Developmentally regulated usage of Physarum DNA replication origins. EMBO Rep *4*, 474-478.

Marilley, M., and Buongiorno-Nardelli, M. (1984). Relationship between the organization of DNA loop domains and of replicons in the eukaryotic genome. Adv Exp Med Biol *179*, 163-168.

Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc Natl Acad Sci U S A *78*, 7634-7638.

Martin, R.M., and Cardoso, M.C. (2010). Chromatin condensation modulates access and binding of nuclear proteins. FASEB J *24*, 1066-1072.

Masai, H. (2013). A personal reflection on the replicon theory: from R1 plasmid to replication timing regulation in human cells. J Mol Biol *425*, 4663-4672.

Masai, H., Matsumoto, S., You, Z., Yoshizawa-Sugata, N., and Oda, M. (2010). Eukaryotic chromosome DNA replication: where, when, and how? Annu Rev Biochem *79*, 89-130.

Masui, S., Nakatake, Y., Toyooka, Y., Shimosato, D., Yagi, R., Takahashi, K., Okochi, H., Okuda, A., Matoba, R., Sharov, A.A., *et al.* (2007). Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. Nat Cell Biol *9*, 625-635.

Matson, J.P., Dumitru, R., Coryell, P., Baxley, R.M., Chen, W., Twaroski, K., Webber, B.R., Tolar, J., Bielinsky, A.K., Purvis, J.E., *et al.* (2017). Rapid DNA replication origin licensing protects stem cell pluripotency. Elife *6*.

Maya-Mendoza, A., Olivares-Chauvet, P., Shaw, A., and Jackson, D.A. (2010). S phase progression in human cells is dictated by the genetic continuity of DNA foci. PLoS Genet *6*, e1000900.

Maya-Mendoza, A., Petermann, E., Gillespie, D.A., Caldecott, K.W., and Jackson, D.A. (2007). Chk1 regulates the density of active replication origins during the vertebrate S phase. EMBO J *26*, 2719-2731.

Mayer, R., Brero, A., von Hase, J., Schroeder, T., Cremer, T., and Dietzel, S. (2005). Common themes and cell type specific variations of higher order chromatin arrangements in the mouse. BMC Cell Biol *6*, 44.

McCarroll, R.M., and Fangman, W.L. (1988). Time of replication of yeast centromeres and telomeres. Cell *54*, 505-513.

Meilinger, D., Fellinger, K., Bultmann, S., Rothbauer, U., Bonapace, I.M., Klinkert, W.E., Spada, F., and Leonhardt, H. (2009). Np95 interacts with de novo DNA methyltransferases, Dnmt3a and Dnmt3b, and mediates epigenetic silencing of the viral CMV promoter in embryonic stem cells. EMBO Rep *10*, 1259-1264.

Meshorer, E., and Misteli, T. (2006). Chromatin in pluripotent embryonic stem cells and differentiation. Nat Rev Mol Cell Biol *7*, 540-546.

Meshorer, E., Yellajoshula, D., George, E., Scambler, P.J., Brown, D.T., and Misteli, T. (2006). Hyperdynamic plasticity of chromatin proteins in pluripotent embryonic stem cells. Dev Cell *10*, 105-116.

Michaelis, C., Ciosk, R., and Nasmyth, K. (1997). Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. Cell *91*, 35-45.

Miller, K.M., Rog, O., and Cooper, J.P. (2006). Semi-conservative DNA replication through telomeres requires Taz1. Nature *440*, 824-828.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*, 631-642.

Miyabe, I., Kunkel, T.A., and Carr, A.M. (2011). The major roles of DNA polymerases epsilon and delta at the eukaryotic replication fork are evolutionarily conserved. PLoS Genet *7*, e1002407.

Mizushima, T., Takahashi, N., and Stillman, B. (2000). Cdc6p modulates the structure and DNA binding activity of the origin recognition complex in vitro. Genes Dev *14*, 1631-1641.

Montagnoli, A., Valsasina, B., Brotherton, D., Troiani, S., Rainoldi, S., Tenca, P., Molinari, A., and Santocanale, C. (2006). Identification of Mcm2 phosphorylation sites by S-phase-regulating kinases. J Biol Chem *281*, 10281-10290.

Moreno, A., Carrington, J.T., Albergante, L., Al Mamun, M., Haagensen, E.J., Komseli, E.S., Gorgoulis, V.G., Newman, T.J., and Blow, J.J. (2016). Unreplicated DNA remaining from unperturbed S phases passes through mitosis for resolution in daughter cells. Proc Natl Acad Sci U S A *113*, E5757-5764.

Mouse Genome Sequencing, C., Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., *et al.* (2002). Initial sequencing and comparative analysis of the mouse genome. Nature *420*, 520-562.

Moyzis, R.K., Buckingham, J.M., Cram, L.S., Dani, M., Deaven, L.L., Jones, M.D., Meyne, J., Ratliff, R.L., and Wu, J.R. (1988). A highly conserved repetitive DNA sequence, (TTAGGG)n, present at the telomeres of human chromosomes. Proc Natl Acad Sci U S A *85*, 6622-6626.

Muto, M., Kanari, Y., Kubo, E., Takabe, T., Kurihara, T., Fujimori, A., and Tatsumi, K. (2002). Targeted disruption of Np95 gene renders murine embryonic stem cells hypersensitive to DNA damaging agents and DNA replication blocks. J Biol Chem *277*, 34549-34555.

Nakamura, H., Morita, T., and Sato, C. (1986). Structural organizations of replicon domains during DNA synthetic phase in the mammalian nucleus. Exp Cell Res *165*, 291-297.

Nakayasu, H., and Berezney, R. (1989). Mapping replicational sites in the eucaryotic cell nucleus. J Cell Biol *108*, 1-11.

Nan, X., Ng, H.H., Johnson, C.A., Laherty, C.D., Turner, B.M., Eisenman, R.N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature *393*, 386-389.

Nasmyth, K., and Haering, C.H. (2009). Cohesin: its roles and mechanisms. Annu Rev Genet *43*, 525-558.

Natale, F., Rapp, A., Yu, W., Maiser, A., Harz, H., Scholl, A., Grulich, S., Anton, T., Horl, D., Chen, W., *et al.* (2017). Identification of the elementary structural units of the DNA damage response. Nat Commun *8*, 15760.

Nativio, R., Wendt, K.S., Ito, Y., Huddleston, J.E., Uribe-Lewis, S., Woodfine, K., Krueger, C., Reik, W., Peters, J.M., and Murrell, A. (2009). Cohesin is required for higher-order chromatin conformation at the imprinted IGF2-H19 locus. PLoS Genet *5*, e1000739.

Newport, J., and Kirschner, M. (1982). A major developmental transition in early Xenopus embryos: I. characterization and timing of cellular changes at the midblastula stage. Cell *30*, 675-686.

Newport, J.W., and Kirschner, M.W. (1984). Regulation of the cell cycle during early Xenopus development. Cell *37*, 731-742.

Niwa, H. (2007). How is pluripotency determined and maintained? Development *134*, 635-646.

Nordman, J., and Orr-Weaver, T.L. (2012). Regulation of DNA replication during development. Development *139*, 455-464.

Norio, P., Kosiyatrakul, S., Yang, Q., Guan, Z., Brown, N.M., Thomas, S., Riblet, R., and Schildkraut, C.L. (2005). Progressive activation of DNA replication initiation in large domains of the immunoglobulin heavy chain locus during B cell development. Mol Cell *20*, 575-587.

O'Keefe, R.T., Henderson, S.C., and Spector, D.L. (1992). Dynamic organization of DNA replication in mammalian cell nuclei: spatially and temporally defined replication of chromosome-specific alpha-satellite DNA sequences. J Cell Biol *116*, 1095-1110.

Oehlmann, M., Score, A.J., and Blow, J.J. (2004). The role of Cdc6 in ensuring complete genome licensing and S phase checkpoint activation. J Cell Biol *165*, 181-190.

Okamoto, I., Otte, A.P., Allis, C.D., Reinberg, D., and Heard, E. (2004). Epigenetic dynamics of imprinted X inactivation during early mouse development. Science *303*, 644-649.

Olovnikov, A.M. (1973). A theory of marginotomy. The incomplete copying of template margin in enzymic synthesis of polynucleotides and biological significance of the phenomenon. J Theor Biol *41*, 181-190.

Pak, D.T., Pflumm, M., Chesnokov, I., Huang, D.W., Kellum, R., Marr, J., Romanowski, P., and Botchan, M.R. (1997). Association of the origin recognition complex with heterochromatin and HP1 in higher eukaryotes. Cell *91*, 311-323.

Palzkill, T.G., and Newlon, C.S. (1988). A yeast replication origin consists of multiple copies of a small conserved sequence. Cell *53*, 441-450.

Panning, M.M., and Gilbert, D.M. (2005). Spatio-temporal organization of DNA replication in murine embryonic stem, primary, and immortalized cells. J Cell Biochem *95*, 74-82.

Papait, R., Pistore, C., Negri, D., Pecoraro, D., Cantarini, L., and Bonapace, I.M. (2007). Np95 is implicated in pericentromeric heterochromatin replication and in major satellite silencing. Mol Biol Cell *18*, 1098-1106.

Parelho, V., Hadjur, S., Spivakov, M., Leleu, M., Sauer, S., Gregson, H.C., Jarmuz, A., Canzonetta, C., Webster, Z., Nesterova, T., *et al.* (2008). Cohesins functionally associate with CTCF on mammalian chromosome arms. Cell *132*, 422-433.

Pauklin, S., Pedersen, R.A., and Vallier, L. (2011). Mouse pluripotent stem cells at a glance. J Cell Sci *124*, 3727-3732.

Payer, B., and Lee, J.T. (2008). X chromosome dosage compensation: how mammals keep the balance. Annu Rev Genet *42*, 733-772.

Perry, P., Sauer, S., Billon, N., Richardson, W.D., Spivakov, M., Warnes, G., Livesey, F.J., Merkenschlager, M., Fisher, A.G., and Azuara, V. (2004). A dynamic switch in the replication timing of key regulator genes in embryonic stem cells upon neural induction. Cell Cycle *3*, 1645-1650.

Peters, A.H., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schofer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A*., et al.* (2001). Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell *107*, 323-337.

Phillips-Cremins, J.E., Sauria, M.E., Sanyal, A., Gerasimova, T.I., Lajoie, B.R., Bell, J.S., Ong, C.T., Hookway, T.A., Guo, C., Sun, Y*., et al.* (2013). Architectural protein subclasses shape 3D organization of genomes during lineage commitment. Cell *153*, 1281-1295.

Pidoux, A.L., and Allshire, R.C. (2005). The role of heterochromatin in centromere function. Philos Trans R Soc Lond B Biol Sci *360*, 569-579.

Pohl, T.J., Brewer, B.J., and Raghuraman, M.K. (2012). Functional centromeres determine the activation time of pericentric origins of DNA replication in Saccharomyces cerevisiae. PLoS Genet *8*, e1002677.

Pope, B.D., Hiratani, I., and Gilbert, D.M. (2010). Domain-wide regulation of DNA replication timing during mammalian development. Chromosome Res *18*, 127-136.

Pope, B.D., Ryba, T., Dileep, V., Yue, F., Wu, W., Denas, O., Vera, D.L., Wang, Y., Hansen, R.S., Canfield, T.K*., et al.* (2014). Topologically associating domains are stable units of replication-timing regulation. Nature *515*, 402-405.

Prasanth, S.G., Shen, Z., Prasanth, K.V., and Stillman, B. (2010). Human origin recognition complex is essential for HP1 binding to chromatin and heterochromatin organization. Proc Natl Acad Sci U S A *107*, 15093-15098.

Quintana-Murci, L., and Fellous, M. (2001). The Human Y Chromosome: The Biological Role of a "Functional Wasteland". J Biomed Biotechnol *1*, 18-24.

Quintana-Murci, L., Krausz, C., and McElreavey, K. (2001). The human Y chromosome: function, evolution and disease. Forensic Sci Int *118*, 169-181.

Quivy, J.P., Roche, D., Kirschner, D., Tagami, H., Nakatani, Y., and Almouzni, G. (2004). A CAF-1 dependent pool of HP1 during heterochromatin duplication. EMBO J *23*, 3516-3526.

Reinhart, M., and Cardoso, M.C. (2017). A journey through the microscopic ages of DNA replication. Protoplasma *254*, 1151-1162.

Rhind, N. (2006). DNA replication timing: random thoughts about origin firing. Nat Cell Biol *8*, 1313-1316.

Rhind, N., and Gilbert, D.M. (2013). DNA replication timing. Cold Spring Harb Perspect Biol *5*, a010132.

Rivera-Mulia, J.C., Buckley, Q., Sasaki, T., Zimmerman, J., Didier, R.A., Nazor, K., Loring, J.F., Lian, Z., Weissman, S., Robins, A.J*., et al.* (2015). Dynamic changes in replication timing and gene expression during lineage specification of human pluripotent stem cells. Genome Res *25*, 1091-1103.

Robertson, K.D. (2005). DNA methylation and human disease. Nat Rev Genet *6*, 597-610.

Rottach, A., Frauer, C., Pichler, G., Bonapace, I.M., Spada, F., and Leonhardt, H. (2010). The multi-domain protein Np95 connects DNA methylation and histone modification. Nucleic Acids Res *38*, 1796-1804.

Rowles, A., Chong, J.P., Brown, L., Howell, M., Evan, G.I., and Blow, J.J. (1996). Interaction between the origin recognition complex and the replication licensing system in Xenopus. Cell *87*, 287-296.

Rowntree, R.K., and Lee, J.T. (2006). Mapping of DNA replication origins to noncoding genes of the X-inactivation center. Mol Cell Biol *26*, 3707-3717.

Ryba, T., Hiratani, I., Lu, J., Itoh, M., Kulik, M., Zhang, J., Schulz, T.C., Robins, A.J., Dalton, S., and Gilbert, D.M. (2010). Evolutionarily conserved replication timing profiles predict long-range chromatin interactions and distinguish closely related cell types. Genome Res *20*, 761-770.

Ryba, T., Hiratani, I., Sasaki, T., Battaglia, D., Kulik, M., Zhang, J., Dalton, S., and Gilbert, D.M. (2011). Replication timing: a fingerprint for cell identity and pluripotency. PLoS Comput Biol *7*, e1002225.

Sado, T., and Sakaguchi, T. (2013). Species-specific differences in X chromosome inactivation in mammals. Reproduction *146*, R131-139.

Sadoni, N., Cardoso, M.C., Stelzer, E.H., Leonhardt, H., and Zink, D. (2004). Stable chromosomal units determine the spatial and temporal organization of DNA replication. J Cell Sci *117*, 5353-5365.

Sansam, C.G., Goins, D., Siefert, J.C., Clowdus, E.A., and Sansam, C.L. (2015). Cyclin-dependent kinase regulates the length of S phase through TICRR/TRESLIN phosphorylation. Genes Dev *29*, 555-566.

Sansoni, V., Casas-Delucchi, C.S., Rajan, M., Schmidt, A., Bonisch, C., Thomae, A.W., Staege, M.S., Hake, S.B., Cardoso, M.C., and Imhof, A. (2014). The histone variant H2A.Bbd is enriched at sites of DNA synthesis. Nucleic Acids Res *42*, 6405-6420.

Sasaki, T., Sawado, T., Yamaguchi, M., and Shinomiya, T. (1999). Specification of regions of DNA replication initiation during embryogenesis in the 65-kilobase DNApolalpha-dE2F locus of Drosophila melanogaster. Mol Cell Biol *19*, 547-555.

Saxonov, S., Berg, P., and Brutlag, D.L. (2006). A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. Proc Natl Acad Sci U S A *103*, 1412-1417.

Schermelleh, L., Haemmer, A., Spada, F., Rosing, N., Meilinger, D., Rothbauer, U., Cardoso, M.C., and Leonhardt, H. (2007). Dynamics of Dnmt1 interaction with the replication machinery and its role in postreplicative maintenance of DNA methylation. Nucleic Acids Res *35*, 4301-4312.

Schwaiger, M., Kohler, H., Oakeley, E.J., Stadler, M.B., and Schubeler, D. (2010). Heterochromatin protein 1 (HP1) modulates replication timing of the Drosophila genome. Genome Res *20*, 771-780.

Selig, S., Ariel, M., Goitein, R., Marcus, M., and Cedar, H. (1988). Regulation of mouse satellite DNA replication time. EMBO J *7*, 419-426.

Sfeir, A., Kosiyatrakul, S.T., Hockemeyer, D., MacRae, S.L., Karlseder, J., Schildkraut, C.L., and de Lange, T. (2009). Mammalian telomeres resemble fragile sites and require TRF1 for efficient replication. Cell *138*, 90-103.

Shapiro, J.A., and von Sternberg, R. (2005). Why repetitive DNA is essential to genome function. Biol Rev Camb Philos Soc *80*, 227-250.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K.*, et al.* (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. Nature *450*, 908-912.

Shermoen, A.W., McCleland, M.L., and O'Farrell, P.H. (2010). Developmental control of late replication and S phase length. Curr Biol *20*, 2067-2077.

Shibata, E., Kiran, M., Shibata, Y., Singh, S., Kiran, S., and Dutta, A. (2016). Two subunits of human ORC are dispensable for DNA replication and proliferation. Elife *5*.

Shin, J.H., Grabowski, B., Kasiviswanathan, R., Bell, S.D., and Kelman, Z. (2003). Regulation of minichromosome maintenance helicase activity by Cdc6. J Biol Chem *278*, 38059-38067.

Sinclair, A.H., Berta, P., Palmer, M.S., Hawkins, J.R., Griffiths, B.L., Smith, M.J., Foster, J.W., Frischauf, A.M., Lovell-Badge, R., and Goodfellow, P.N. (1990). A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. Nature *346*, 240-244.

Singh, N.P., Madabhushi, S.R., Srivastava, S., Senthilkumar, R., Neeraja, C., Khosla, S., and Mishra, R.K. (2011). Epigenetic profile of the euchromatic region of human Y chromosome. Nucleic Acids Res *39*, 3594-3606.

Singleton, M.R., Dillingham, M.S., and Wigley, D.B. (2007). Structure and mechanism of helicases and nucleic acid translocases. Annu Rev Biochem *76*, 23-50.

Sjogren, C., and Strom, L. (2010). S-phase and DNA damage activated establishment of sister chromatid cohesion--importance for DNA repair. Exp Cell Res *316*, 1445-1453.

Solovei, I., and Joffe, B. (2010). Inverted nuclear architecture and its development during differentiation of mouse rod photoreceptor cells: a new model to study nuclear architecture. Genetika *46*, 1159-1163.

Solovei, I., Kreysing, M., Lanctot, C., Kosem, S., Peichl, L., Cremer, T., Guck, J., and Joffe, B. (2009). Nuclear architecture of rod photoreceptor cells adapts to vision in mammalian evolution. Cell *137*, 356-368.

Solovei, I., Thanisch, K., and Feodorova, Y. (2016). How to rule the nucleus: divide et impera. Curr Opin Cell Biol *40*, 47-59.

Song, J., Lafont, A., Chen, J., Wu, F.M., Shirahige, K., and Rankin, S. (2012). Cohesin acetylation promotes sister chromatid cohesion only in association with the replication machinery. J Biol Chem *287*, 34325-34336.

Song, J., Rechkoblit, O., Bestor, T.H., and Patel, D.J. (2011). Structure of DNMT1-DNA complex reveals a role for autoinhibition in maintenance DNA methylation. Science *331*, 1036-1040.

Soufi, A., and Dalton, S. (2016). Cycling through developmental decisions: how cell cycle dynamics control pluripotency, differentiation and reprogramming. Development *143*, 4301-4311.

Speck, C., and Stillman, B. (2007). Cdc6 ATPase activity regulates ORC x Cdc6 stability and the selection of specific DNA sequences as origins of DNA replication. J Biol Chem *282*, 11705-11714.

Splinter, E., Heath, H., Kooren, J., Palstra, R.J., Klous, P., Grosveld, F., Galjart, N., and de Laat, W. (2006). CTCF mediates long-range chromatin looping and local histone modification in the beta-globin locus. Genes Dev *20*, 2349-2354.

Sporbert, A., Domaing, P., Leonhardt, H., and Cardoso, M.C. (2005). PCNA acts as a stationary loading platform for transiently interacting Okazaki fragment maturation proteins. Nucleic Acids Res *33*, 3521-3528.

Sporbert, A., Gahl, A., Ankerhold, R., Leonhardt, H., and Cardoso, M.C. (2002). DNA polymerase clamp shows little turnover at established replication sites but sequential de novo assembly at adjacent origin clusters. Mol Cell *10*, 1355-1365.

Stanojcic, S., Sollelis, L., Kuk, N., Crobu, L., Balard, Y., Schwob, E., Bastien, P., Pages, M., and Sterkers, Y. (2016). Single-molecule analysis of DNA replication reveals novel features in the divergent eukaryotes Leishmania and Trypanosoma brucei versus mammalian cells. Sci Rep *6*, 23142.

Stead, E., White, J., Faast, R., Conn, S., Goldstone, S., Rathjen, J., Dhingra, U., Rathjen, P., Walker, D., and Dalton, S. (2002). Pluripotent cell division cycles are driven by ectopic Cdk2, cyclin A/E and E2F activities. Oncogene *21*, 8320-8333.

Stevens, T.J., Lando, D., Basu, S., Atkinson, L.P., Cao, Y., Lee, S.F., Leeb, M., Wohlfahrt, K.J., Boucher, W., O'Shaughnessy-Kirwan, A.*, et al.* (2017). 3D structures of individual mammalian genomes studied by single-cell Hi-C. Nature *544*, 59-64.

Stewart, S.A., and Weinberg, R.A. (2006). Telomeres: cancer to human aging. Annu Rev Cell Dev Biol *22*, 531-557.

Suter, B., Tong, A., Chang, M., Yu, L., Brown, G.W., Boone, C., and Rine, J. (2004). The origin recognition complex links replication, sister chromatid cohesion and transcriptional silencing in Saccharomyces cerevisiae. Genetics *167*, 579-591.

Takagi, N., and Sasaki, M. (1975). Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. Nature *256*, 640-642.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. Cell *126*, 663-676.

Takebayashi, S., Sugimura, K., Saito, T., Sato, C., Fukushima, Y., Taguchi, H., and Okumura, K. (2005). Regulation of replication at the R/G chromosomal band boundary and pericentromeric heterochromatin of mammalian cells. Exp Cell Res *304*, 162-174.

Tamaru, H., and Selker, E.U. (2001). A histone H3 methyltransferase controls DNA methylation in Neurospora crassa. Nature *414*, 277-283.

Tanaka, S., Nakato, R., Katou, Y., Shirahige, K., and Araki, H. (2011). Origin association of Sld3, Sld7, and Cdc45 proteins is a key step for determination of origin-firing timing. Curr Biol *21*, 2055-2063.

Tariq, M., Saze, H., Probst, A.V., Lichota, J., Habu, Y., and Paszkowski, J. (2003). Erasure of CpG methylation in Arabidopsis alters patterns of histone H3 methylation in heterochromatin. Proc Natl Acad Sci U S A *100*, 8823-8827.

Techer, H., Koundrioukoff, S., Azar, D., Wilhelm, T., Carignon, S., Brison, O., Debatisse, M., and Le Tallec, B. (2013). Replication dynamics: biases and robustness of DNA fiber analysis. J Mol Biol *425*, 4845-4855.

Terret, M.E., Sherwood, R., Rahman, S., Qin, J., and Jallepalli, P.V. (2009). Cohesin acetylation speeds the replication fork. Nature *462*, 231-234.

Thomsen, N.D., and Berger, J.M. (2009). Running in reverse: the structural basis for translocation polarity in hexameric helicases. Cell *139*, 523-534.

Ticau, S., Friedman, L.J., Ivica, N.A., Gelles, J., and Bell, S.P. (2015). Single-molecule studies of origin licensing reveal mechanisms ensuring bidirectional helicase loading. Cell *161*, 513-525.

Tolhuis, B., Palstra, R.J., Splinter, E., Grosveld, F., and de Laat, W. (2002). Looping and interaction between hypersensitive sites in the active beta-globin locus. Mol Cell *10*, 1453-1465.

Tsumura, A., Hayakawa, T., Kumaki, Y., Takebayashi, S., Sakaue, M., Matsuoka, C., Shimotohno, K., Ishikawa, F., Li, E., Ueda, H.R.*, et al.* (2006). Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. Genes Cells *11*, 805-814.

Uemura, T., Kubo, E., Kanari, Y., Ikemura, T., Tatsumi, K., and Muto, M. (2000). Temporal and spatial localization of novel nuclear protein NP95 in mitotic and meiotic cells. Cell Struct Funct *25*, 149-159.

Unoki, M., Nishidate, T., and Nakamura, Y. (2004). ICBP90, an E2F-1 target, recruits HDAC1 and binds to methyl-CpG through its SRA domain. Oncogene *23*, 7601-7610.

Usdin, K., House, N.C., and Freudenreich, C.H. (2015). Repeat instability during DNA repair: Insights from model systems. Crit Rev Biochem Mol Biol *50*, 142-167.

Van der Aa, N., Cheng, J., Mateiu, L., Zamani Esteki, M., Kumar, P., Dimitriadou, E., Vanneste, E., Moreau, Y., Vermeesch, J.R., and Voet, T. (2013). Genome-wide copy number profiling of single cells in S-phase reveals DNA-replication domains. Nucleic Acids Res *41*, e66.

van Dierendonck, J.H., Keyzer, R., van de Velde, C.J., and Cornelisse, C.J. (1989). Subdivision of S-phase by analysis of nuclear 5-bromodeoxyuridine staining patterns. Cytometry *10*, 143-150.

van Driel, R., Fransz, P.F., and Verschure, P.J. (2003). The eukaryotic genome: a system regulated at different hierarchical levels. J Cell Sci *116*, 4067-4075.

Vermeulen, M., Eberl, H.C., Matarese, F., Marks, H., Denissov, S., Butter, F., Lee, K.K., Olsen, J.V., Hyman, A.A., Stunnenberg, H.G.*, et al.* (2010). Quantitative interaction proteomics and genome-wide profiling of epigenetic histone marks and their readers. Cell *142*, 967-980.

Vig, B.K. (1995). The centromere: kinetochore complex. Southeast Asian J Trop Med Public Health *26 Suppl 1*, 68-76.

Vogelauer, M., Rubbi, L., Lucas, I., Brewer, B.J., and Grunstein, M. (2002). Histone acetylation regulates the time of replication origin firing. Mol Cell *10*, 1223-1233.

Wade, P.A., Gegonne, A., Jones, P.L., Ballestar, E., Aubry, F., and Wolffe, A.P. (1999). Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. Nat Genet *23*, 62-66.

Walter, J., and Newport, J. (2000). Initiation of eukaryotic DNA replication: origin unwinding and sequential chromatin association of Cdc45, RPA, and DNA polymerase alpha. Mol Cell *5*, 617-627.

Walter, J., and Newport, J.W. (1997). Regulation of replicon size in Xenopus egg extracts. Science *275*, 993-995.

Weidtkamp-Peters, S., Rahn, H.P., Cardoso, M.C., and Hemmerich, P. (2006). Replication of centromeric heterochromatin in mouse fibroblasts takes place in early, middle, and late S phase. Histochem Cell Biol *125*, 91-102.

Weiler, K.S., and Wakimoto, B.T. (1995). Heterochromatin and gene expression in Drosophila. Annu Rev Genet *29*, 577-605.

Wendt, K.S., Yoshida, K., Itoh, T., Bando, M., Koch, B., Schirghuber, E., Tsutsumi, S., Nagae, G., Ishihara, K., Mishiro, T.*, et al.* (2008). Cohesin mediates transcriptional insulation by CCCTC-binding factor. Nature *451*, 796-801.

White, J., and Dalton, S. (2005). Cell cycle control of embryonic stem cells. Stem Cell Rev *1*, 131-138.

Wijchers, P.J., Geeven, G., Eyres, M., Bergsma, A.J., Janssen, M., Verstegen, M., Zhu, Y., Schell, Y., Vermeulen, C., de Wit, E.*, et al.* (2015). Characterization and dynamics of pericentromere-associated domains in mice. Genome Res *25*, 958-969.

Wijgerde, M., Grosveld, F., and Fraser, P. (1995). Transcription complex stability and chromatin dynamics in vivo. Nature *377*, 209-213.

Wohlschlegel, J.A., Dhar, S.K., Prokhorova, T.A., Dutta, A., and Walter, J.C. (2002). Xenopus Mcm10 binds to origins of DNA replication after Mcm2-7 and stimulates origin binding of Cdc45. Mol Cell *9*, 233-240.

Wong, P.G., Winter, S.L., Zaika, E., Cao, T.V., Oguz, U., Koomen, J.M., Hamlin, J.L., and Alexandrow, M.G. (2011). Cdc45 limits replicon usage from a low density of preRCs in mammalian cells. PLoS One *6*, e17533.

Woodfine, K., Fiegler, H., Beare, D.M., Collins, J.E., McCann, O.T., Young, B.D., Debernardi, S., Mott, R., Dunham, I., and Carter, N.P. (2004). Replication timing of the human genome. Hum Mol Genet *13*, 191-202.

Wright, W.E., Tesmer, V.M., Huffman, K.E., Levene, S.D., and Shay, J.W. (1997). Normal human chromosomes have long G-rich telomeric overhangs at one end. Genes Dev *11*, 2801-2809.

Wright, W.E., Tesmer, V.M., Liao, M.L., and Shay, J.W. (1999). Normal human telomeres are not late replicating. Exp Cell Res *251*, 492-499.

Yamashita, M., Hori, Y., Shinomiya, T., Obuse, C., Tsurimoto, T., Yoshikawa, H., and Shirahige, K. (1997). The efficiency and timing of initiation of replication of multiple replicons of Saccharomyces cerevisiae chromosome VI. Genes Cells *2*, 655-665.

Yu, C., Gan, H., Han, J., Zhou, Z.X., Jia, S., Chabes, A., Farrugia, G., Ordog, T., and Zhang, Z. (2014). Strand-specific analysis shows protein binding at replication forks and PCNA unloading from lagging strands when forks stall. Mol Cell *56*, 551-563.

Zellner, E., Herrmann, T., Schulz, C., and Grummt, F. (2007). Site-specific interaction of the murine pre-replicative complex with origin DNA: assembly and disassembly during cell cycle transit and differentiation. Nucleic Acids Res *35*, 6701-6713.

Zeman, M.K., and Cimprich, K.A. (2014). Causes and consequences of replication stress. Nat Cell Biol *16*, 2-9.

Zhang, P., Hastert, F.D., Ludwig, A.K., Breitwieser, K., Hofstatter, M., and Cardoso, M.C. (2017a). DNA base flipping analytical pipeline. Biology Methods and Protocols *2*, 1-12.

Zhang, P., Rausch, C., Hastert, F.D., Boneva, B., Filatova, A., Patil, S.J., Nuber, U.A., Gao, Y., Zhao, X., and Cardoso, M.C. (2017b). Methyl-CpG binding domain protein 1 regulates localization and activity of Tet1 in a CXXC3 domain-dependent manner. Nucleic Acids Res *45*, 7118-7136.

Zhang, Y., Ng, H.H., Erdjument-Bromage, H., Tempst, P., Bird, A., and Reinberg, D. (1999). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. Genes Dev *13*, 1924-1935.

Zhong, Y., Nellimoottil, T., Peace, J.M., Knott, S.R., Villwock, S.K., Yee, J.M., Jancuska, J.M., Rege, S., Tecklenburg, M., Sclafani, R.A.*, et al.* (2013). The level of origin firing inversely affects the rate of replication fork progression. J Cell Biol *201*, 373-383.

Zhou, J.C., Janska, A., Goswami, P., Renault, L., Abid Ali, F., Kotecha, A., Diffley, J.F.X., and Costa, A. (2017). CMG-Pol epsilon dynamics suggests a mechanism for the establishment of leading-strand synthesis in the eukaryotic replisome. Proc Natl Acad Sci U S A *114*, 4141-4146.

Zou, Y., Gryaznov, S.M., Shay, J.W., Wright, W.E., and Cornforth, M.N. (2004). Asynchronous replication timing of telomeres at opposite arms of mammalian chromosomes. Proc Natl Acad Sci U S A *101*, 12928-12933.

LIST OF ABBREVIATIONS

(m)ESC(s)	(murine) embryonic stem cell(s)
3D-SIM	3D structured illumination microscopy
ACS	Autonomous Consensus Sequence
ARS	Autonomously Replicating Sequence
BrdU	5-bromo-2'-deoxyuridine
CDC45	*Cell Division Cycle 45*
CDC6	*cell division control protein 6*
CDK	*cyclin-dependent kinase*
CDT1	*Chromatin licensing and DNA replication factor 1*
CldU	5-chloro-2'-deoxyurdine
CMG	*complex consisting of CDC45, MCM2-7 and GINS*
CpG	CG dinucleotide
CTCF	*CCCTC-binding factor*
CTD	C-terminal domain
DAPI	4',6-diamidino-2-phenylindole
DDK	*Dfb4-dependent kinase*
DHFR	*dihydrofolate reductase*
DNTM1/3A/3B	*DNA methyltransferases 1/3A/3B*
dsDNA	double-stranded DNA
EdU	5-ethynyl-2'-deoxyuridine
FA	fork asymmetrie
FACS	fluorescence activated cell sorting
FISH	fluorescence in-situ hybridization
FRAP	fluorescence recovery after photobleaching
GINS	*`Go-Ichi-Ni-San`; complex consisting of Sld5, Psf1, Psf3 and Psf3*
GS	genome size
HP1	*heterochromatin protein 1*
ICM	inner cell mass
IdU	5-iodo-2'-deoxyurdine
IOD	inter-origin distance
KD	knockdown
LAD	lamina-associated domain
LEF(s)	loop extrusion factor(s)

MaSat	major satellite repeats
MBT	mid-blastula transition
MCM2-7	*minichromosome maintenance complex, subunits 2 - 7*
MEF	mouse embryonic fibroblast
MiSat	minor satellite repeats
nanoRF(i)	nano replication focus/replication foci
NPC(s)	neuronal progenitor cell(s)
NTD	N-terminal domain
ORC1-6	*Origin Recognition Complex, subunits 1-6*
PCNA	*proliferating cell nuclear antigen*
periHC	pericentromeric heterochromatin
pMCM	phospho-MCM
pol α	*polymerases alpha/primase*
pol δ	*polymerases delta*
pol ε	*polymerases epsilon*
preRC	pre-replication complex
pWF	pseudo wide-field
RC	replication complex
RD	replication domain
RF(i)	replication focus/replication foci
RFC	*replication factor C*
RFS	replication fork speed
RPA	*replication protein A*
RT	replication timing
SSB	single-stranded binding protein
ssDNA	single-stranded DNA
SU	subunit

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

[redacted]

DECLARATION – EHRENWÖRTLICHE ERKLÄRUNG

Ehrenwörtliche Erklärung:

Ich erkläre hiermit ehrenwörtlich, dass ich die vorliegende Arbeit entsprechend den Regeln guter wissenschaftlicher Praxis selbstständig und ohne unzulässige Hilfe Dritter angefertigt habe.

Sämtliche aus fremden Quellen direkt oder indirekt übernommenen Gedanken sowie sämtliche von Anderen direkt oder indirekt übernommenen Daten, Techniken und Materialien sind als solche kenntlich gemacht. Die Arbeit wurde bisher bei keiner anderen Hochschule zu Prüfungszwecken eingereicht.

Darmstadt, den 19. September 2018

.....
Patrick Weber

CURRICULUM VITAE

Personal Information

Name:	Weber, Patrick
Date of Birth:	18.12.1988
Place of Birth:	Mannheim
Address:	Schnittspahnstr. 10, 64287 Darmstadt, Germany

Doctoral studies

10/2013 – 10/2018	PhD Thesis at the Technical University Darmstadt Department of Biology, Cell Biology & Epigenetics Prof. Dr. M. Cristina Cardoso Topic: "DNA replication dynamics in embryonic stem cells"

Academic Education

2011 – 2013	M. Sc. at the Technical University Darmstadt Prof. Dr. M Cristina Cardoso Topic: “DNA replication dynamics in murine embryonic stem cells” Final Grade: 1,2
2008 – 2011	B. Sc. at the Technical University Darmstadt Prof. Dr. Markus Löbrich Topic: "Radiation-induced changes in gene expression relating to proliferation and differentiation of fibroblasts: Influence of exogenous TGF-beta1 in relation to intraoperative radiotherapy (IORT)" Final Grade: 2,1

Publications

Under revision

Weber, *P., Rausch, C., Scholl, A., Cardoso, M. C. (2018). Repli-FISH: Application of 3D-(Immuno)-FISH for the study of DNA replication timing of genetic repeat elements. OBM Genetics*

Published

Becker, A., Allmann, L., Hofstatter, M., Casa, V., **Weber, P**., Lehmkuhl, A., Herce, H. D. and Cardoso, M. C. (2013). Direct homo- and hetero-interactions of MeCP2 and MBD2. PLoS One 8: e53730.